

The SAGE Encyclopedia of
**THEORY IN SCIENCE,
TECHNOLOGY, ENGINEERING,
AND MATHEMATICS**

James M. Mattingly
Editor



2

The SAGE Encyclopedia of
**THEORY IN SCIENCE,
TECHNOLOGY,
ENGINEERING, AND
MATHEMATICS**

The SAGE Encyclopedia of
**THEORY IN SCIENCE,
TECHNOLOGY,
ENGINEERING, AND
MATHEMATICS**

1

Edited by

James Mattingly

Georgetown University

 **SAGE** reference

Los Angeles | London | New Delhi | Singapore | Washington DC | Melbourne



FOR INFORMATION:

SAGE Publications, Inc.
2455 Teller Road
Thousand Oaks, California 91320
E-mail: order@sagepub.com

SAGE Publications Ltd.
1 Oliver's Yard
55 City Road
London, EC1Y 1SP
United Kingdom

SAGE Publications India Pvt. Ltd.
B 1/I 1 Mohan Cooperative Industrial Area
Mathura Road, New Delhi 110 044
India

SAGE Publications Asia-Pacific Pte. Ltd.
18 Cross Street #10-10/11/12
China Square Central
Singapore 048423

Copyright © 2023 by SAGE Publications, Inc.

All rights reserved. Except as permitted by U.S. copyright law, no part of this work may be reproduced or distributed in any form or by any means, or stored in a database or retrieval system, without permission in writing from the publisher.

All third party trademarks referenced or depicted herein are included solely for the purpose of illustration and are the property of their respective owners. Reference to these trademarks in no way indicates any relationship with, or endorsement by, the trademark owner.

Printed in the United States of America

ISBN: 9781483347721

Acquisitions Editor: Kaitlin Ciarmiello
Development Editor: Sanford Robinson
Production Editor: Astha Jaiswal
Copy Editor: Hurix Digital
Typesetter: Hurix Digital
Proofreaders: Dennis Web, Ellen Brink, Theresa Kay
and Wendy Jo Dymond
Indexer: Sheila Hill
Cover Designer: Candice Harman
Marketing Manager: Gabrielle Perretta

This book is printed on acid-free paper.

22 23 24 25 26 10 9 8 7 6 5 4 3 2 1

Contents

Volume 1

List of Entries	<i>vii</i>
Reader's Guide	<i>xi</i>
About the Editor	<i>xv</i>
List of Contributors	<i>xvii</i>
Introduction	<i>xxi</i>

Entries

A	1	G	299
B	57	H	343
C	113	I	363
D	209	J	423
E	229	K	429
F	279	L	439

Volume 2

List of Entries	<i>vii</i>
Reader's Guide	<i>xi</i>

Entries

M	501	T	893
N	603	U	931
p	629	V	935
R	767	W	945
S	799		

Appendix: Resource Guide	957
--------------------------	-----

Index	961
-------	-----

List of Entries

Abduction
Abstract Knowledge
Abstraction
Accuracy
Actor-Network Theory
Ad Hoc Hypothesis
Analysis
Applicability of Mathematics in Science
Artificial Intelligence
Assumption
Astronomy. *See* Cosmology
Authority
Axiom Schema
Axiomatic Theory

Belief Revision
Big Data
Biochemistry, 19th Century
Biochemistry, 20th Century
Biochemistry, Contemporary
Bioinformatics
Biology, Evolutionary
Biophysics, 19th Century
Biophysics, 20th Century
Biophysics, Contemporary
Biopsychosocial Model
Biostatistics

Category Theory
Cell Theory
Chemistry
Church–Turing Thesis
Climate Science
Cognitive Science
Communication Theory, Technical Overview
Complementary Medicine
Completeness
Complex Systems
Computer Science. *See* Cybernetics, 20th Century;
Software Engineering; Big Data

Concept
Conceptual Analysis
Conceptual Blending
Consilience
Cosmology
Cybernetics, 20th Century

Data and Phenomena
Data Mining. *See* Big Data
Data Models
Deduction
Developmental Systems Theory

Ecology. *See* Environmental Studies; Natural
Resources
Empiricism. *See* Epistemology; Evidence;
Experiment, Theory of
Engineering, Contemporary
Engineering, History of
Environmental Studies
Epistemology
Evidence
Evolutionary Psychology
Experiment, Theory of
Explanation

Fact Versus Theory
Falsifiability
Formal Sciences
Framework

Game Theory
Generalization
Genetic Drift
Geology
Geometry, Classical
Geometry, Non-Euclidean
Geophysics. *See* Geology
Germ Theory
Gödel’s Incompleteness Theorems

Health Care Science
Hypothesis Testing
Hypothetico-Deductivism

Induction
Infectious Disease Studies
Inference
Inference to the Best Explanation
Inferentialism
Informatics
Information Theory
Information Theory, Historical Background
Instrumentalism
Interpretation
Intuition
Intuitionism in Logic and Mathematics

Justification

Kinetic (Molecular) Theory
Knowledge

Laws, Scientific
Laws of Nature
Life Sciences, Contemporary
Linguistic Frameworks
Linguistics, Contemporary
Linguistics, Historical
Logic, Formal and Informal
Logic, Inductive
Logic and Language
Logical Theory

Mathematics, Antiquity
Mathematics, Enlightenment
Mathematics, Middle Ages
Mathematics, 19th Century
Mathematics, Renaissance
Mathematics, 20th Century
Measurement
Medicine, Antiquity
Medicine, Contemporary
Medicine, Enlightenment
Medicine, Middle Ages
Medicine, 19th Century
Medicine, Renaissance
Medicine, 20th Century. *See* Medicine,
Contemporary; Medicine, 19th Century
Mental Models
Metaphysics

Metatheory
Meteorology. *See* Climate Science
Modeling

Nanotechnology
Natural Resources
Natural Selection
Neuroscience
Number Theory. *See* Mathematics, 19th Century;
Mathematics, 20th Century

Paraconsistency
Paradigm
Paradoxes
Perturbation Theory
Phenomenalism
Philosophy of Mind
Philosophy of Science
Physical Theory
Physics, Antiquity
Physics, Contemporary
Physics, Enlightenment
Physics, Gravitation Theory
Physics, Middle Ages
Physics, 19th Century
Physics, Quantum Theory
Physics, Renaissance
Physics, Solid-State
Physics, Thermodynamics
Physics, 20th Century
Plate Tectonics
Pragmatism
Prediction
Pseudoscience
Punctuated Equilibrium

Rational Mechanics
Rationality
Realism in Mathematics
Realism in Science. *See* Scientific
Realism
Reasoning
Received View of Theories. *See* Theory,
Syntactic Conception
Refutation
Relative Consistency

Scientific Realism
Scientific Revolutions
Semantics (Introduction to Theory)

Semantics (Scientific and Empirical)	Taxonomy
Set Theory	Theory Change
Social Construction of Scientific Knowledge	Theory Construction
Software Engineering	Theory Structure
Soundness	Theories, Semantic Conception of
Speculation	Theories, Syntactic Conception of
Statistical Inference, Bayesian	Thought Experiments, Scientific and Philosophical
Statistical Inference, Frequentist	Truth
Statistics	Understanding
Statistics, Completeness in	Values in Science
Structuralism (Introduction to Theory)	Vitalism
Syntax (Introduction to Theory)	Warrant
Syntax (Scientific and Empirical)	Worldview
Systems Science	

Reader's Guide

General Theory of Science

Abduction
Abstract Knowledge
Abstraction
Accuracy
Ad Hoc Hypothesis
Analysis
Authority
Belief Revision
Church–Turing Thesis
Communication Theory, Technical Overview
Completeness
Concept
Conceptual Analysis
Conceptual Blending
Consilience
Epistemology
Evidence
Experiment, Theory of
Explanation
Fact Versus Theory
Falsifiability
Formal Sciences
Generalization
Hypothesis Testing
Hypothetico-Deductivism
Induction
Inference
Inference to the Best Explanation
Inferentialism
Instrumentalism
Interpretation
Intuition
Justification
Knowledge
Mental Models
Metaphysics
Metatheory
Modeling
Paraconsistency

Paradigm
Paradoxes
Phenomenalism
Philosophy of Mind
Philosophy of Science
Physical Theory
Pragmatism
Prediction
Pseudoscience
Rationality
Realism in Mathematics
Reasoning
Scientific Realism
Scientific Revolutions
Social Construction of Scientific Knowledge
Speculation
Statistics
Taxonomy
Theory Change
Thought Experiments, Scientific and Philosophical
Truth
Understanding
Values in Science

Nature and Structure of Theories

Axiom Schema
Axiomatic Theory
Data Models
Framework
Geometry, Classical
Geometry, Non-Euclidean
Gödel's Incompleteness Theorems
Instrumentalism
Intuitionism in Logic and Mathematics
Laws, Scientific
Laws of Nature
Linguistic Frameworks
Modeling
Paradigm
Philosophy of Science

Rational Mechanics
 Received View of Theories
 Relative Consistency
 Set Theory
 Theories, Semantic Conception of
 Theories, Syntactic Conception of
 Theory Construction
 Theory Structure

Formal Disciplines

Formal Sciences
 Game Theory

Computer Science

Artificial Intelligence
 Big Data
 Bioinformatics
 Biostatistics
 Church–Turing Thesis
 Complex Systems
 Cybernetics, 20th Century
 Informatics
 Information Theory
 Information Theory, Historical Background
 Software Engineering
 Statistical Inference, Bayesian
 Statistical Inference, Frequentist
 Statistics
 Statistics, Completeness in
 Systems Science

Logic and Mathematics

Abduction
 Abstraction
 Analysis
 Axiom Schema
 Axiomatic Theory
 Category Theory
 Church–Turing Thesis
 Communication Theory, Technical
 Overview
 Concept
 Cybernetics, 20th Century
 Deduction
 Generalization
 Geometry, Classical
 Geometry, Non-Euclidean
 Gödel's Incompleteness Theorems
 Induction
 Inference
 Intuitionism in Logic and Mathematics

Justification
 Knowledge
 Linguistics, Contemporary
 Linguistics, Historical
 Logic, Formal and Informal
 Logic, Inductive
 Logic and Language
 Logical Theory
 Mathematics, 19th Century
 Mathematics, 20th Century
 Mathematics, Antiquity
 Mathematics, Enlightenment
 Mathematics, Middle Ages
 Mathematics, Renaissance
 Paraconsistency
 Paradoxes
 Perturbation Theory
 Rational Mechanics
 Realism in Mathematics
 Reasoning
 Refutation
 Relative Consistency
 Semantics (Introduction to Theory)
 Semantics (Scientific and Empirical)
 Set Theory
 Soundness
 Statistics
 Syntax (Introduction to Theory)
 Syntax (Scientific and Empirical)
 Understanding

Empirical Disciplines

Physical Theory

Biological Science

Biochemistry, 19th Century
 Biochemistry, 20th Century
 Biochemistry, Contemporary
 Bioinformatics
 Biology, Evolutionary
 Biophysics, 19th Century
 Biophysics, 20th Century
 Biophysics, Contemporary
 Biostatistics
 Cell Theory
 Complementary Medicine
 Developmental Systems Theory
 Environmental Studies
 Evolutionary Psychology
 Game Theory
 Genetic Drift

Germ Theory
 Health Care Science
 Infectious Disease Studies
 Life Sciences, Contemporary
 Medicine, 19th Century
 Medicine, 20th Century
 Medicine, Antiquity
 Medicine, Contemporary
 Medicine, Enlightenment
 Medicine, Middle Ages
 Medicine, Renaissance
 Natural Selection
 Neuroscience
 Punctuated Equilibrium
 Vitalism

Chemistry

Biochemistry, 19th Century
 Biochemistry, 20th Century
 Biochemistry, Contemporary
 Chemistry

Cognitive Sciences

Artificial Intelligence
 Belief Revision
 Big Data
 Biopsychosocial Model
 Cognitive Science
 Conceptual Blending
 Evolutionary Psychology
 Linguistic Frameworks
 Linguistics, Contemporary
 Linguistics, Historical
 Mental Models
 Neuroscience
 Philosophy of Mind
 Semantics (Introduction to Theory)
 Semantics (Scientific and Empirical)
 Syntax (Introduction to Theory)
 Syntax (Scientific and Empirical)

Earth and Space Sciences

Climate Science
 Cosmology
 Environmental Studies
 Geology
 Natural Resources
 Plate Tectonics

Engineering

Big Data

Bioinformatics
 Communication Theory, Technical
 Overview
 Complex Systems
 Cybernetics, 20th Century
 Engineering, Contemporary
 Engineering, History of
 Informatics
 Information Theory
 Modeling
 Perturbation Theory
 Software Engineering
 Statistics

Physics

Biophysics, 19th Century
 Biophysics, 20th Century
 Biophysics, Contemporary
 Complex Systems
 Cosmology
 Environmental Studies
 Geophysics
 Information Theory
 Kinetic (Molecular) Theory
 Physical Theory
 Physics, 19th Century
 Physics, 20th Century
 Physics, Antiquity
 Physics, Contemporary
 Physics, Enlightenment
 Physics, Gravitational Theory
 Physics, Middle Ages
 Physics, Quantum Theory
 Physics, Renaissance
 Physics, Solid-State
 Physics, Thermodynamics
 Plate Tectonics
 Rational Mechanics
 Systems Science

Technology

Artificial Intelligence
 Cybernetics, 20th Century
 Engineering, Contemporary
 Engineering, History of Health Care
 Science
 Informatics
 Modeling
 Nanotechnology
 Software Engineering

About the Editor

James Mattingly is Associate Professor of Philosophy at Georgetown University. Originally from the Silicon Valley in California, he studied Great Books at St. John's College in Annapolis, Maryland, and physics at University of California, Santa Cruz. He then returned to a study of the history and philosophy of science at Indiana University, where he received his PhD in 2002. Shortly before that, he was appointed assistant professor in the Georgetown University Department of Philosophy, where he has been ever since. His research is primarily in philosophy of science. He spends his efforts there about equally between general issues involving conceptual change in the sciences, the epistemology of science, the nature of scientific theories, scientific experiments, and scientific explanation on the one hand, and issues more

specific to philosophy of physics on the other, including quantum gravity, general relativity, black holes and singularities, gauge theories, thermodynamics, and electrodynamics. He also has research interests in early modern philosophy, the foundations of logic and mathematics, and the history of logical empiricism and other movements that attempted to come to grips with the profound conceptual reorientation made necessary by the revolutionary changes in science at the turn of the 20th century. He has written and lectured extensively on scientific theories, epistemology of science, the nature of scientific experimentation, and the foundations of physics. He is the author of *Information and Experimental Knowledge*, published in 2021 by The University of Chicago Press.

List of Contributors

Marcus P. Adams
*The State University of New
York at Albany*

Syed Feroj Ahmed
*CSIR–Indian Institute of
Chemical Biology*

Olayinka Akanle
University of Ibadan, Nigeria

Mattea Andreoletti
*Vita-Salute San Raffaele
University, Italy*

André Ariew
University of Missouri

Anthony C. Atkinson
London School of Economics

Warren J. Bareiss
*University of South Carolina
Upstate*

Magdalena Bielenia-Grajewska
University of Gdansk

Robert C. Bishop
Wheaton College

Amy Sue Bix
Iowa State University

David F. Bjorklund
Florida Atlantic University

Gina Bloodworth
Salisbury University

Alisa Bokulich
Boston University

Derek Bolton
King's College London

Nora Mills Boyd
Siena College

Claudia Bucciferro
*Rochester Institute of
Technology*

Sandro Caparrini
Independent Scholar

Gordon S. Carlson
Fort Hays State University

Michael J. Carter
*California State University,
Northridge*

Mark Casaupin
Independent Scholar

Jose A. Chamizo
*Autonomous National
University of Mexico*

Ian C. Clift
*Biomedical Associates, Granger,
Indiana*

Justin Corfield
Independent Scholar

Robert P. Crease
Stony Brook University

Benjamin L. Curtis
Independent Scholar

Douglas J. Dallier
Winona State University

Gesiel da Silva
Independent Scholar

Michael Davis
Illinois Institute of Technology

Gregory W. Dawes
University of Otago

David DeIulius
Bethany College

Sami Dhungana
Cal Poly, San Luis Obispo

Vikdor Dorfler
University of Strathclyde

Amitabh Vikram Dwivedi
*Shri Mata Vaishno Devi
University*

Taner Edis
Truman State University

Yavuz Ercil
Baskent University

Nicolas Fillion
Simon Fraser University

Thora L. Fitzpatrick
Independent Scholar

Courtney Vail Fletcher
University of Portland

Nathan E. Fosse
Harvard University

John Mark Froiland
Purdue University

Michelle Gibbons
University of New Hampshire

Peter Gildenhuys
Lafayette College

Ido Golding
*University of Illinois at
Urbana–Champaign*

João Carlos Graça
University of Lisbon, Portugal

Jeremy Gray
Open University

Janice D. Hamlet
Northern Illinois University

- | | | |
|---|--|---|
| Jason Hannan
<i>University of Winnipeg</i> | Eric Malczewski
<i>Virginia Tech</i> | Naomi Jeffery Petersen
<i>Central Washington University</i> |
| Sam Haraway
<i>University of California, Davis</i> | Carlos Eduardo Maldonado
<i>Universidad El Bosque</i> | J. Brian Pitts
<i>University of Lincoln</i> |
| Chris Haufe
<i>Independent Scholar</i> | Luca Mari
<i>Università Càattaneo LIUC</i> | Anthony Puddephatt
<i>Lakehead University</i> |
| Lenadra H. Hernandez
<i>National University</i> | Owen J. E. Maroney
<i>Independent Scholar</i> | Nicholas Ramsey
<i>University of California, Los Angeles</i> |
| Nicholaos Jones
<i>The University of Alabama in Huntsville</i> | Craig E. Mattson
<i>Trinity Christian College</i> | Ole Ravn
<i>Aalborg University</i> |
| Premalatha Karupiah
<i>University Sains Malaysia</i> | Mandy M. McBroom
<i>Independent Scholar</i> | Leslie J. Reynard
<i>Washburn University</i> |
| Louis H. Kauffman
<i>University of Illinois at Chicago</i> | Trudy Mercadal
<i>Florida Atlantic University</i> | Joe Roussos
<i>Institute for Futures Studies</i> |
| Neha Katrapal
<i>O. P. Jinda Global University</i> | Daniele Molinini
<i>University of Lisbon</i> | Joseph C. Rumenapp
<i>St. John's University</i> |
| Benjamin W. Kelly
<i>Nipissing University</i> | Katie Moss
<i>Independent Scholar</i> | Alexander Ruser
<i>University of Agder</i> |
| Samuel Kimpton-Nye
<i>The University of Bristol</i> | Roman Murawski
<i>Adam Mickiewicz University</i> | Robert Rynasiewicz
<i>Johns Hopkins University</i> |
| Samson Olowo Kolawole
<i>Nigeria Police Academy</i> | Sarah E. Naramore
<i>Northwest Missouri State University</i> | Henrik Skaug Sætra
<i>Østfold University College</i> |
| Helge Kragh
<i>University of Copenhagen</i> | Mariam Orkodashvili
<i>Vanderbilt University</i> | Rodolfo Sarsfield
<i>Autonomous University of Queretaro</i> |
| Gregory Landini
<i>University of Iowa</i> | Asil Ali Özdögrü
<i>Üsküdar University</i> | Ulysses Sengupta
<i>Manchester School of Architecture</i> |
| Marc Lange
<i>University of North Carolina at Chapel Hill</i> | Rupa Pal
<i>University of Calcutta</i> | James A. Shapiro
<i>University of Chicago</i> |
| Richard Leo Lanigan
<i>International Communicology Institute</i> | William M. Peaster
<i>Independent Scholar</i> | Mark. D. Sherry
<i>University of Toledo</i> |
| Alex Levine
<i>University of South Florida</i> | Judy Pelham
<i>York University</i> | Robert Sinclair
<i>Soka University</i> |
| Pablo Lorenzano
<i>UNQ/CONICET</i> | Pablo A. Pellegrini
<i>CONICET/National University of Quilmes</i> | Mo E. Snyder
<i>Acadia University</i> |
| Sebastian Lutz
<i>Uppsala University, Sweden</i> | Michelle E. Pence
<i>The University of Texas of the Permian Basin</i> | Aris Spanos
<i>Virginia Tech</i> |
| | Nuria Perez-Perez
<i>University of Pompeu Fabra</i> | Jan Sprenger
<i>University of Turin</i> |

Walter Scott Stepanenko
Independent Scholar

Richard Teague
Johns Hopkins University

Rafael E. Testa
University of Campinas

Klaus G. Troitzsch
University of Koblenz (Retired)

Riva Tukachinsky
Chapman University

Murat Pasa Uysal
Özyegin University

Timothy I. Vandiver
Independent Scholar

Ioannis Votsis
New College of the Humanities

Uanina Wellmann
Leuphana Universität Lüneburg

Frank Welz
Innsbruck University

David C. Witherington
University of New Mexico

Susan J. Wurtzburg
Independent Scholar

Qinghua Yang
University of Miami

Qingjiang Yao
Lamar University

Dimitry Zakharov
Independent Scholar

Scott Zimmer
Independent Scholar

Introduction

The word “theory” has fascinating origins. In ancient Greek, the primary meaning of the word “θεωρία” (theory) seems to have been “looking at, viewing, beholding”.¹ It also referred to ambassadors sent out to observe the oracle, or, on other occasions, the Olympic games. A variation of the word, *theorem*, can refer to the results of some science or technical discipline (what we might think of as engineering), and this is the word that Euclid uses to refer to the things that he proves in the *Elements*, our first systematic geometry book. In those cases, the theorems, or the things observed using the discipline, were tightly linked to the act of showing and seeing, to making plain. Today we might refer to, say, number theory and some of the things that we discover or prove or explain using number theory as its theorems. While “theory” is often associated with abstraction, with complication, with abstruseness it is grounded in the concrete, observable, repeatable results of the crafts, mathematics, and science.

The word “theory” is also, ironically, tightly linked equally to certainty and to pure speculation. On the one hand, in the popular imagination, theory seems sometimes associated to, say, the quantum theory, or the theory of relativity, two disciplines that have achieved unprecedented accuracy and experimental confirmation in the century since they were first formulated. On the other hand, it is common to hear, in response to a claim about the origin of various species on the earth, that natural selection is *only* a theory. Generally, people who study scientific theories insist, and rightly so, that speculation ungrounded in prior disciplinary work and without strong observational support does not merit being called “theory.” Theories are generally well-supported explanatory

systems that organize the understanding of some discipline or sector of it. While novel theories are propounded in order to try to explain puzzling or novel data and are not accepted until they have themselves been tested, when used to make assertions about the world and what we can expect to see, a theory typically has already been stringently tested and found to be observationally adequate.

Why an Encyclopedia of Theory, and Why STEM?

A collection of articles about scientific theory in its various aspects seems worthwhile for a number of reasons. First and foremost our scientific theories are the vehicles of what we know where the results and understanding of various disciplines are codified so that they can be transmitted, taught to new members of the discipline, and integrated with the results of other disciplines. They are, relatedly, tools for understanding and controlling the phenomena studied in a discipline. This is enough to make the theories themselves of significant and enduring interest. But in addition to the multitude of individual theories in the various STEM disciplines, there is also the matter of theory as itself a topic worthy of study.

My own academic discipline is neither science, nor technology, nor engineering, nor mathematics. Instead, it is philosophy of science. Philosophers of science do make contributions, sometimes, to science (or technology, or engineering, or mathematics) but that is not what their main focus is. Their main focus is to understand science and related disciplines as *themselves* objects of study. They ask: How does physics relate to chemistry? What do biologists mean when they use the expression “species”? Do experiments in social science give secure knowledge? Beyond such discipline specific questions though philosophers of science are

1. H. G. Liddell & R. Scott (Eds.). (1945). *An intermediate Greek-English lexicon*. Oxford University Press.

interested in big questions about the nature of scientific inquiry, and one of the main foci of philosophy of science has long been the theories that are produced and used in the sciences. What kind of thing is a theory? Is it a collection of true remarks about some field? Is it a calculus for predicting things? Is it a vehicle for understanding the subject matter of some discipline? Or is it something else altogether? These are important concerns to philosophers. But philosophers' concerns are not always the concerns of scientists, and are rarely those of students. We are often seen as out of touch with the real, grounded, day to day understanding of things and we are seen as nit-picking to little effect. Happily, however, philosophers are not the only ones who ask and answer such questions. So too do sociologists of science, scientific practitioners themselves, engineers, mathematicians, logicians, and so on. These have joined in with philosophers to contribute to this encyclopedia. What you will find across the many entries in this encyclopedia is a broad diversity of disciplinary background, but it is a diversity that displays unity of purpose if not of method or disciplinary assumptions. These many perspectives highlight just some of the important ways of understanding theory in philosophy and the sciences and make it clear that there are a number of roads into an understanding of theory.

As noted earlier, this encyclopedia is designed to support the study of theory in itself as well as the study of specific theories. The reasons for this are many. First, the scope of the encyclopedia allows for highlighting the vast range of topics that fall under the concept and also highlight key developments of theory along with key uses in various STEM disciplines. Next, putting things together this way can help to normalize the proper sense of "theory" as it is used in the sciences, and allow for nuanced discussion of the limits and scope of the claims made in the sciences without falling into bimodal habits of excessive deference or thoughtless dismissal. The concreteness of the specific disciplines' theories and the variety of tools and techniques that arise in one and find application in another provides an interesting contrast. This contrast can help to illustrate the fruitfulness over time of feedback between study of theory and development of theory as an object of use in STEM disciplines, and transfer of innovations in theory

between disciplines. Finally, this encyclopedia, covering as it does such a range of topics displays the importance of interactions and feedback between theory as it used and theory as it is . . . well . . . theorized.

The encyclopedia is an attempt to shed some light on the nature of theories in the STEM disciplines. I have talked about reasons to focus on theory. But why focus on STEM? First, it may not seem to be a good category in some respects, either as an academic umbrella term or as a unified object of study. The classification of STEM fields under an umbrella term that joins the abstract pursuits of modern sciences like quantum theory and mathematics together with the concrete and practical practices of technological production and engineering may seem like a novelty thought up by ambitious administrators trying to figure out where best to direct institutional efforts and funding at fields that are likely to return measurable fruits. While there is some truth to that, it is almost precisely what Robert Boyle and folks like him were trying to do when they instituted the practice of modern science. Out of fractured and individualized practices and closely guarded techniques and knowledge developed in secret, Boyle attempted to produce an open society of inquiry, grounded in observation and reason, dedicated to improving human lives.

In any case, STEM is a term of current use, and moreover can be useful here as a demarcation criterion between a number of ways "theory" can be employed. In some academic disciplines, theory has more to do with providing an interpretation of various kinds of activity or features of the world so that some new phenomena can be revealed. Maybe an analysis of various patterns of speech in a community can be seen to reveal important background assumptions of that community, and sometimes "theorizing" is a way of capturing the activity involved in making it possible to be aware of those assumptions. This use of "theory" has its own life and vast disciplinary context, one that is importantly different from what will be found in these pages. Here the use of "theory," although broad, is constrained in different ways from much of what is found in other disciplines. It would be silly to say that one use is correct and the other not, and it would also be silly to suppose that there is nothing to be learned about one way of

deploying the concept theory from studying other ways, for there will be important points of contact. However, STEM is a natural domain of application for one way of thinking about the concept, and there is utility for students in fields broadly under this umbrella, or who are trying to learn about topics related to those fields, to have a ready and broadly unified resource for orienting themselves to the way theory is used and developed in those fields. And as an aside it is the one that hews most closely to my own disciplinary competence.

The Organization and Rationale of the Encyclopedia

My main aims for the encyclopedia are two. On the one hand, I want users to find here examples of the various technical and conceptual components of what it is to be a theory, in general terms and not merely as these theories appear in and are used by various STEM focused disciplines as a way of organizing, grasping, making use of, critiquing, and expanding our knowledge of some subject matter. To reflect that part of the aim included here are a number of entries that are concerned with theory itself, in some way detached from any particular subject matter concerning which one might construct or employ a theory. On the other hand, I want users to see the way that various STEM disciplines theorize about their subject matters, both how they make use of existing understanding of what theories are and how they function and how the individual disciplines, by attempting to make theories adequate to the peculiarities of their subject matter, fruitfully extend the very meaning of what it is to be a theory. While these subject matter entries will contain framing material about the theory of their subject matter, their main task will be to illuminate what role theory plays in the discipline and what insights into the nature of theory follows from that illumination.

The entries for the *Encyclopedia of Theory in STEM* range widely. There are 173 entries comprising nearly 625,000 words of text in entries ranging from 1,500 to 7,000 words. The entries have been grouped into 13 categories: there are four broad categories with several further categories grouped under two of them. The categories and their respective entries form the Reader's

Guide, which follows the List of Entries in the front matter of the encyclopedia. The categories are also listed below. I have chosen to group according to discipline, widely understood, as the most perspicuous way of navigating the topics of the encyclopedia. There is inevitable categorical overlap in a number of disciplines, and some inevitable awkward fits given our scope constraints, and the fact that some of the disciplines considered here are generally thought to be situated in non-STEM-covering disciplines, but seemed necessary to include for reasons of coherence.

The first division in the structure of the encyclopedia is between *general* matters having to do with science and theory and *specific* matters having to do with particular disciplines within STEM.

I. General Theory of Science: Here you will find entries having to do with the nature of science, including its authoritativeness, the way its claims are tested, in short the basic conceptual features of the scientific enterprise as a whole.

II. Nature and Structure of Theories: Here the discussion remains general and at a high level of abstraction but the focus is more specifically on the nature of scientific theory. What kind of thing is a theory? How is it connected to other things we do in the sciences? What have theorists about theories learned over the last century from science and from analysis of science, and what has science itself learned from such theorizing.

Turning to the more specific disciplines we see the second division, between disciplines more on the empirical side of things and those on the formal side. Here are found discussions both of matters having to do specifically with the day-to-day features of these disciplines and of general issues that frame and clarify their scope and limits. While these entries give accessible overviews of the subject matter of the disciplines, those overviews are no substitute for extensive study of the disciplines. Instead just enough is presented to make it possible to understand the role of theory in these disciplines.

III. Formal Disciplines: These disciplines are less grounded in empirical data and more focused on concepts. But that is not to say that empirical matters can be entirely neglected. Even in logic and

mathematics, but especially in computer science questions of resources and the role of complexity make it impossible to fully detach oneself from facts about the world. Even so entries here are more focused on structural features of theories, in addition to specific disciplinary matters. This topic is further divided into:

1. Computer Science
2. Logic and Mathematics

IV. Empirical Disciplines: The empirical disciplines as well are not monolithically empirical. Constraints from the nature of mathematics and logic have an impact on the kinds of theory that are possible here. But these entries are more directed to specific developments in observationally testable domains as well as to the connection between theoretical and observational features. This topic is further divided into:

1. Biological Science
2. Chemistry
3. Cognitive Sciences
4. Earth and Space Sciences
5. Engineering
6. Physics
7. Technology

Acknowledgments

This project would not exist without Geoffrey Golson, who initially approached me about trying to sort out what an encyclopedia in this neighborhood would look like. His strong leadership was crucially important in the early stages. As work progressed, the value of the editorial acumen and diligence of Diana Axelsen would be hard to overstate. This project has been long in the making and has had to be reconceived a number of times. Andrew Boney and Sanford Robinson were wonderful collaborators in those efforts of revision, and the encyclopedia benefited at every turn from their contributions. Tamara Tanso provided invaluable assistance to keep me on track and up-to-date with the entries, and the behind-the-scenes support from Sage staff especially made bearable the many quotidian tasks required for such endeavors. Christian Golden's contributions resonate through all these pages, and I thank him profoundly for his good humor, his understanding, and his powerful work ethic. Finally, Natalia makes everything possible and worth doing.

James Mattingly

A

ABDUCTION

Abduction is a form of reasoning distinct from deductive and inductive reasoning. It has often been referred to as the *commonsense* approach to reasoning because it is used in everyday life as well as in scientific argumentation. It has also been called *inference to the best explanation* because people reason from something that has been observed to an explanation for it. The philosopher Charles Sanders Peirce (1839–1914) first described this form of reasoning as *guessing* because he noted that you must leap to a possible explanation without knowing with certainty whether or not it is true. Abductive reasoning is a critical notion in philosophy, including the philosophy of science, as well as that of history, law, and many other disciplines. It is also a basic function of the human mind in everyday life.

Abduction in Everyday Life

Abductive reasoning is used in daily life as a commonsense approach to explaining anomalies. One of the most common examples given is that of walking out of a house to find that the grass in the lawn is wet. Because the grass is wet, one may infer that it had recently rained. If it had rained, then the fact that the grass is wet would be nothing out of the ordinary. That it had rained explains the fact that the grass is wet.

Another such case used to illustrate the process of abduction is that of a faulty light fixture. If one

comes home and flips a light switch to *on*, but the light does not turn on, then one may infer that the power is out or the breaker switch has been flipped. Suppose that, although the light is still off, one sees that the clock on the microwave is on and appears to be functioning correctly. If the power were off in the whole house, then this contradiction would not occur because the microwave clock would be off as well. Now, given this new evidence, one may infer a simpler explanation to account for the light being off, namely that the light bulb has burned out and needs to be changed. In this way, abduction is the process of inferring explanations to account for the evidence or facts that are observed. The explanation can never be proven but rather only inferred.

As a generic tool of human reasoning, abduction has been of interest to philosophers because, unlike deductive reasoning, abduction is not a formalized system of proofs. In fact, the subjectivity required to decide which explanation is best has been one of the difficulties in applying abduction to science and other domains.

Abduction in Use

Abduction is commonly used in philosophical inquiry, in belief revision, in academic and professional disciplines such as medicine and anthropology, and in work-related reasoning. Philosophers of science also draw on abduction as *inference to the best explanation* to be one of the primary tools in scientific realism. Abduction is a component part of human reasoning. There have been attempts to

formalize abductive reasoning to meet the needs of particular jobs or sectors of society. Some decisions require a type of validation, or rules to decide which possible explanation for an observed effect is best. On the most basic level, abduction has no such rules. It is minimalistic in this aspect and allows for subjectivity in determining the best explanation, often leading to revising previously held beliefs. For the hypothesis to be a useful one, it should be explanatory, testable, and economical. The simplest explanation is usually expected to be the best, though the subjective view can never escape the possibility for error. Below are several areas of application, and many of the uses of abduction take place among practitioners or clinicians within a particular field.

Logical Reasoning

Logical reasoning has often been split into *deductive* and *inductive* reasoning. The former begins with reasoning from a general premise to an entailed, or necessary, conclusion. There is room for error if the formal rules of logic have been followed. The latter, on the other hand, is reasoning from specific evidence or premises to a probable conclusion. The rules are not formally set out, but evidence can be collected until it sufficiently can support a conclusion.

Abductive reasoning, according to Peirce, provides a third type of reasoning by which one reasons from a conclusion to a possible explanation. This presents some particular difficulties in the realm of formal logic. Formal logic would propose that the process of inferring is deductive, as in the instance that follows:

All bachelors are male.

John is a bachelor.

Therefore, *John is male.*

Formally, this is written as follows where A is the antecedent and C is the consequent:

$A \rightarrow C$

A

—————

$\therefore C$

In this example of *modus ponens*, we see that given the two premises ($A \rightarrow C$ and A), we must conclude that the conclusion (C) is valid. Whether or not it is true is not a concern. Logical validity is established by adhering to certain rules of deduction. Reasoning, then, happens from the given premise to infer a necessary conclusion. Abduction, however, operates in the reverse.

Abduction begins with a given conclusion or consequent, C and then seeks to infer an explanation for C. This is the formal logical fallacy of affirming the consequent because there is no guarantee that by knowing C one can infer A. That is, returning to the example above, that if *John* is male, there is no guarantee that John is a *bachelor*. Thus, abduction is a form of inference that violates formal logic, but is nonetheless critical to human thinking and philosophical reasoning. Formally, this would be written as follows:

$A \rightarrow C$

C

—————

A “Invalid”

While it is clear that abduction and deduction are quite different types of reasoning, the distinction between abduction and induction is less clear. Since inductive reasoning is the process of learning from example, some have proposed that induction is only a case of abduction. That is, inductive reasoning can be reduced to the more primal abductive reasoning. Other philosophers, who view induction as creating a theory to account for facts, have proposed that abduction is actually only a case of inductive reasoning. Peirce’s triad of reasoning can be found to have overlap in actual use. Thus, some have proposed to use the terminology of *explanatory reasoning* to account for the specific process and product of abduction specifically.

Medical Sciences

The medical sciences make use of abduction, especially in clinical practice, through observing the symptoms of a patient and reasoning as to what might have caused the current condition. A

physician may attempt to explain the cause of several symptoms experienced by a patient. Sometimes, the coexistence of several symptoms may complicate the diagnostic process as a doctor seeks to identify the root cause. Usually, the explanation that seems to explain all of the symptoms is preferable to diagnosis of several causes, thus drawing on abductive reasoning's simplicity to pinpoint a probable cause that can then be treated. Diagnosis and clinical evaluation are examples of abduction in medicine.

Law

Legal experts and lawyers often utilize abductive reasoning. Specifically within criminal cases, a jury makes a judgment regarding whether the prosecution's case sufficiently explains the evidence. The burden, then, is on the prosecution to present a plausible abduction that a judge or jury could endorse as a true narrative of events leading to the crime. The defense, on the other hand, needs only to present an alternative abduction or cast reasonable doubt on the prosecution's explanation of the evidence, to secure the verdict of not guilty. Developing a theory that can be inferred from and explain the evidence is critical in jurisprudence.

Statistics and Applied Mathematics

Abduction has been influential in applied mathematics and inferential statistics. Bayesian inferential statistics, for example, has been a useful way to make decisions based on statistical data. Through statistical evidence, one could reject or accept a hypothesis depending on whether the probability is sufficient for a particular cause to have produced the observed effect. The level of probability represents the most likely explanation, and tells the statistician which hypothesis is most statistically likely, though not necessarily the best, explanation for a consequent. This helps to remove some of the subjectivity innate to abduction. It restricts *the best explanation* to that which is *most likely*.

Abduction also has been used to develop artificial intelligence in computer programming. Parameters for abductive processes can be set, which would allow machines to analyze data and make inferences to explain that data. By modeling

artificial intelligence with abduction, machines can mimic human thinking. Abduction allows for the detection of faults in computer systems, as well as belief revision. This is contrasted with deduction, which would require preprogramming all possible data so the machine can abide by the formal rules of logic.

Philosophy of Science

The influence of abduction as an area of study is perhaps most clearly seen in the philosophy of science. Since the emphasis in abductive reasoning is to infer possible explanations, it can be seen that abduction does not lead to proof or certainty. As noted in the logical fallacy above, one cannot determine the antecedent from the consequent with absolute certainty. One may only infer that the antecedent is a probable cause for the consequent because the presence of the antecedent would sufficiently explain the consequent. Abduction, then, has been foundational in the area of scientific reasoning and, more broadly, what counts as science. Biological diversity, for example, can be explained by many theories. Two of the most debated are Charles Darwin's account of natural selection and creationist accounts of a divine maker. Though either of the two could serve as reasons for biological diversity, the majority of modern-day scientists have appealed to abductive reasoning to validate Darwinian evolution as the operating theory. Since one could claim that Darwin's approach sufficiently explains biological diversity, and is more plausible than a divine maker, scientists reason that the best explanation for biological diversity is natural selection. Science, then, allows for uncertain and unobserved possible causes to be taken as operational *truth* or *fact* insofar as they explain the evidence well enough. This necessarily introduces a degree of subjectivity.

The Subjectivity of Abduction

Abduction has been the target of some critique, specifically regarding the inherent subjectivity. For example, the association with abduction as the *inference to the best explanation* would require an interpretive move as to what is considered the *best* explanation for the accepted evidence. Additionally, it has fallen under criticism because it is used as a

form of knowledge discovery, not as scientific verification. Thus, abduction is sometimes seen simply as hypothesis forming, and not as a valid form of logic or hypothesis verification. In epistemology, abduction can never achieve *truth* or *proof* claims but can only propose possible explanations.

A Scientific Example of Abduction

Abduction is a fundamental concept in philosophy of science. Examples of abduction can be found frequently in scientific inferences. However, one example that serves to illustrate the point is the scientific discovery of the earth's orbit and rotation. Through observation of the sun, moon, and stars, one would be tempted to think that the earth is the center of the universe and that celestial bodies revolve around it. Ptolemy's (ca. 127–145 C.E.) geocentric view was certainly influential, and it was accepted as true since it best accounted for the evidence at hand. Surely, this is what the ancients thought and indeed how they lived their lives. They observed the phenomena that the sun rises in the east and sets in the west, and the most clear-cut explanation is that the sun revolves around the earth. Some observers sought to explain that it was not the sun that was moving, but the earth that was rotating. Philolaus, the Pythagorean (470 B.C.E.–385 B.C.E.), for example, attempted to account for the same observance by theorizing that there was actually a non-earth fire-center around which the sun, moon, and planets circulated. The view was quite complicated and largely dismissed. Others had thought it possible that the earth rotated, thereby giving the appearance that the sun rose and set, but these views had failed to explain all of the cosmic observations and left early scientists believing that geocentrism was the best explanation. It was not until Nicolaus Copernicus (1473–1543 C.E.) developed his heliocentric view that the seeming anomalies in the cosmos were better explained.

Prior to Copernicus, no system naturally explained retrograde planetary motion where, at some times of the year, planets would seem to move backward in their orbits, and then forward again. Ptolemy's system needed to appeal to *epicycles*, the idea that planets were on smaller orbits which were themselves on larger orbits around the earth, to explain the qualitative features of

retrograde motion. Thus, the models presented were highly complicated and made little theoretical sense though they were inferred from the observational evidence. When Copernicus developed his version of heliocentrism, he made inferences from his observations that the way to best explain the celestial movements would be if the sun was in the center and all of the planets revolved around it. This would make some planets appear to move backward because the earth would catch up to them in the orbit and bypass them. Copernicus' system also used epicycles, in fact just as many as did Ptolemy's, but only for *quantitative* features of things. The explanation for the observations was then simplified. Surely, the claim that the sun was the center of the solar system and that the earth revolved around it and rotated on its axis was not without its challengers. But through further scientific validation it was commonly accepted, verified, and improved upon. What began with observation of the sky ended with an inference as to what could possibly account for these observations. No one could actually sit on the knees of God, so to speak, and see the universe's workings, but from one perspective the workings of the universe could be inferred. Thus, science, as well as philosophy, relies on the inference to the best explanation, or abduction, to understand what may cause observed effects.

Joseph C. Rumenapp

See also Deduction; Explanation; Fact Versus Theory; Induction; Inference; Inference to the Best Explanation; Theory Construction; Values in Science

Further Readings

- Aliseda, A. (2006). *Abductive reasoning: Logical investigations into discovery and explanation*. Springer.
- Douven, I. (2011, Spring). Abduction. In E. N. Zalta (Ed.), *The Stanford encyclopedia of philosophy*. <http://plato.stanford.edu/archives/spr2011/entries/abduction/>
- Harman, G. (1965). The inference to the best explanation. *Philosophical Review*, 74(1), 88–95.
- Lipton, P. (2004). *Inference to the best explanation* (2nd ed.). Routledge.
- Peirce, C. S. (1931–1935; 1958). *Collected papers of Charles Sanders Peirce* (Vols. 1–6, C. Hartshorne & P. Weiss (Eds.); Vols. 7–8 A. W. Burks (Ed.)). Harvard University Press.

ABSTRACT KNOWLEDGE

The term *abstract* originates from the Latin *abstractus*, from the prefix *ab-*, meaning “from,” “of”; and “*trahere*,” which means “pull,” “drag away.” It refers to the mode of thinking that allows for the isolation of elements. The topic of abstract thinking can be examined from a philosophical, psychological, and even a neurophysiological standpoint. It generally assumes that the individual is capable of mentally deconstructing a whole into its parts, in order to analyze a situation simultaneously from different angles. In this way, the practice of abstract thought leads to abstract knowledge, which builds upon the individual’s ability to find common properties among elements or ideas, to plan and assume hypothetically, and to think and act symbolically.

Abstract Thought and the Human World

Among the disciplines that study abstract knowledge are biology-related sciences. Ever since Charles Darwin’s (1809–1882) theories of natural selection emerged, scientists have tried to identify which physical elements differentiate human beings, and their thinking processes, from those of other animals. Theorists believe that as humans created increasingly complex social models, these led to more conceptually sophisticated forms of thinking and knowledge. In order to survive and thrive, for example, humans had to develop ways of understanding the world from multiple perspectives, to draw inferences about social patterns, and regulate their own thinking and emotions according to normative group standards. Understanding and incorporating societal cues and norms form the trove of abstract knowledge that allows humans to navigate the social world around them. Language and culture, for example, were born from the need to cooperate and create complex thinking models in order to grow. What differentiates humans from other animals, as some scientists propose, are forms of thinking, learning, and cognition engendered by collaborative and communicative interaction. These interactions rely on information constructed from symbols and assumptions, which require the capability of abstract thought.

In time, humans developed cognitive capacities far more powerful than was necessary for a hunting-and-gathering group to cooperate and survive. Darwin believed that the capacity and use of language might have shaped the highly complex human brain. Subsequently, however, both biology and the cognitive sciences long evaded the issue. Some behaviorists perceived abstract thinking simply as normal animal communication rather than as manifestations of intelligence and language. The brain, working mechanically, was capable of self-organizing in order to save and process new input, such as words. Words, however, are inseparably tied to the concepts they represent. As such, they are part of a cognitive process that merges thinking and words into symbolic representation and meaningful combinations, so as to form a learned language. Therefore, language is considered one of the many instances of human abstract knowledge.

Psychologists have also been interested in developing theories of abstract knowledge. Abstract thought is different from concrete thinking, which is based on real, lived experiences. Early in life, individuals come to understand their daily experiences as real and as based upon interaction with concrete objects. In time, they begin to develop abstract thinking, that is, their own ideas or concepts. It is relevant to mention the work of Swiss psychologist Jean Piaget (1896–1980), who posited that abstract thinking actually develops at around 12 years of age. At this point, humans move from concrete thinking to being able to mentally explore abstract or immaterial ideas. Before that, children are more reliant on concrete thinking. An example of concrete thinking would be the case of very young children who believe that if they cover their eyes, they cannot be seen by others, because they, themselves, cannot see. Unless they have some mental impairment, however, adolescents and adults can understand abstract concepts, such as comprehending that the phrase “a bad taste in my mouth” can refer to a feeling that something is false, deceitful, or unfair, rather than the literal feeling of having a disagreeable taste in the mouth. In other words, abstract knowledge makes it possible to comprehend metaphors and similes as symbolic representations or approximations of ideas. Not all metaphors, however, refer to abstract ideas. The phrase “Get

out of here!” may have a real as well as a metaphorical meaning, the latter referring to an expression of incredulity. People can understand the difference because, inevitably, abstract knowledge is often built upon concrete concepts.

Early philosophical concepts of abstract thinking were first developed by the ancient Greek philosopher Aristotle (384–322 B.C.E.), especially in his theory of universals. Abstraction is a mental activity in which a conceptual property is detached from the whole, so that it may be reflected upon and analyzed, separated from the rest. Whereas Aristotle’s teacher Plato (ca. 429–348 B.C.E.) believed that all objects and ideas had an essence or pure form, which could be perceived intuitively by humans, Aristotle posited that universal ideas exist but that these must be founded upon empirical data or observation. In other words, abstract knowledge must be built upon concrete information. In that way, the idea of *tree* comes from the experience and comparison of many trees, from which aggregate an individual may extrapolate a general concept of *tree*, which shares the characteristics of what they all have in common. It is then by comparing multiple objects that a property considered shared by all may be isolated or abstracted, and the isolated object is a universal. The individual then forms an idea of what a tree, which represents all trees, is. This becomes abstract knowledge, based upon exercising abstract thinking. Abstract thinking is not necessarily based on concrete or material experiences but on the capacity of making projections and building conjectures on a purely abstract and intangible plane. In this sense, many also call it *theoretical*, considered separate from material or concrete existence. One example of abstract thinking, for example, is developed in mathematics and physics, as well as in many philosophical concepts that deal with intangible ideas. Others have compared abstract thinking with the metaphorical plane, which implies the ability to represent ideas symbolically.

It is important to highlight that mental representations of real objects are aided by language. Abstract knowledge is based on formal schemas, which are thought units by which knowledge is represented. Schemas make it possible for people to have foresight and think in terms of probabilities. They also allow the people to categorize and

integrate new information. Abstract thinking is fundamental to all human beings because it allows the capacity of induction and deduction, to synthesize information, and to extrapolate what is important to learn from all situations, as well as to compare, arrive at conclusions, and engage in critical thinking. In short, abstract thinking builds up abstract knowledge, which allows individuals to interpret the concrete phenomena they come across. One of the characteristics of higher level abstract thought is the capacity of simultaneously observing myriad details and evaluating various functions, as well as processing several problems, defining priorities, and responding to several tasks. Besides judgment, foresight, and reasoning, abstract thinking is also related to creativity, insight, and mental flexibility in general.

Abstract Thinking and Language

Abstract knowledge relies on language in order to conceptualize intangible concepts, such as honor, freedom, and justice, as opposed to more concrete items, such as table and tree. Abstract language is also necessary in order to relay the ideas behind metaphors, allegories, similes, idioms, and other figures of speech. Concrete language is used, for example, to explain more textual instructions, such as in manuals, medical protocols, or cooking recipes. Abstract language is used to explain more abstract ideas, such as calculus or equations.

Abstract language is also related to relative concepts. There are, of course, many degrees of abstraction. Intellectual development, as noted earlier, is tied to the gradual process of moving from the very concrete to extremely abstract thinking across a wide range of areas. In many ways, both the concrete and the abstract are specific to a particular area. For some experts, familiarity with an area makes the topic, however abstract, increasingly concrete. In this manner, then, abstract concepts can become relatively concrete to some while remaining highly abstract to others. For instance, a mathematician might find concepts of differential equations to be relatively concrete, although they remain abstract to others; an atom may be a concrete entity to a physicist, while it may be considered an abstract concept by nonexperts.

Language is a system of word-based communication that relies on both symbolic and conventional

characteristics and which is complemented by culturally based nonverbal communication and word content. Culture is the framework that mediates social actions. In order to acquire language, humans must move from learning language to developing an internal language, also known as *thinking*. This knowledge, although abstract and internal, is mediated by culture and society. Piaget's theory of cognitive constructivism, for example, views language as an element of symbolic function in society. Language has a semiotic function, a process of articulation between figurative elements and their actions. The figurative elements of language are the signifiers and signified, what the words are and what they mean. Semiotics, a theory of language, is an example of knowledge which uses concrete and abstract ways of thinking, developed by scholars such as Charles Sanders Peirce and Ferdinand de Saussure.

Language and abstract thought, then, are intrinsically related in many ways. Although the ability of children to manage abstract concepts develops as they acquire language, spoken language is not the sole component in abstract thinking. Children who have been born deaf, for example, can understand many abstract concepts even though they may not develop speech language fluently. They can, however, express themselves symbolically by way of signed language and other forms of communication available to them. Abstract knowledge, then, is based on multiple forms of language and communication.

By thinking abstractly, individuals elaborate relationships of meanings, such as similarities, differences, and causalities, with symbolic meaning represented by way of language. There are three general types of thinking related to language: conceptual, judgment, and reasoning. *Conceptual thinking* refers to the representation of an object by way of abstraction and generalization. *Abstraction* provides the selection of the properties of that object which distinguish it from others, and *generalizing* refers to the assignation of common characteristics to all of those objects that share the same characteristics. Generalizations allow for different levels of depth and directions of thought that can be taken. For every problem, there are some specific directions and levels that may be more useful than others. A computer, however, can be programmed to generalize concepts to levels of

abstraction and proliferation that presumably would exceed human capabilities, such as in chess-playing programs. Abstract thought is the skill, and abstract and concrete knowledge provide the tools that enable an individual to identify which directions and levels are more useful, or which direction is the most interesting. It is in these decisions that human individuality often shines forth and in which human imagination often beats the computer program.

Judgment allows the elaboration of premises that will determine the truth or falseness of a conclusion and reasoning allows the elaboration of analysis by establishing the relationship between one or more judgments or premises. Reasoning includes deductive and inductive thinking, in where deduction refers to conclusions that are arrived by way of going from the general to the specific, and induction goes from the specific to the general. An example of deductive thinking might be the formulation of the following: All humans are intelligent. All those present here are human. Therefore, all those present here are intelligent. An example of inductive thinking might be as follows: Copper expands with heat. Bronze expands with heat. Both are metals. Therefore, all metals expand with heat.

One of the most explicit theories relating language and thought is the theory of linguistic relativity developed by Benjamin Whorf (1897–1941). Whorf argued that the language structure used by an individual originates from their particular conception of the world. Different people see the world in a unique way and employ language accordingly. Therefore, an individual's language expresses a person's determined concept of the world. Many scholars believe that the cognition, thinking processes, and abstract knowledge of an individual are inevitably affected by the particular language structure of each person. For Whorf, the range of vocabulary that a person learns influences the range of activities that the person may conceive related to a particular area. It may also influence the range of creativity and abstract ideas that the person may reach. Whorf's hypothesis has two main versions. One states that language directly impacts thought. Another, more nuanced, states that language affects thought when a particular task depends directly on the properties of a specific language system. There is very little in the way of

proof, to date, on the first version of Whorf's hypothesis. Language, however, is generally considered as one of the dominant elements in abstract thinking and knowledge. The abstracting process that is ever-present in language can be considered as representative of this intellectual activity.

Another important theorist who has written on language and abstract knowledge is the Soviet psychologist Lev Vygotsky (1896–1934). Vygotsky's theories on the relationship between thinking, knowledge, and language have had great impact in contemporary psychology, especially in the arena of evolutionary psychology. Vygotsky argues that thinking and language, high-level mental activities, have different genetic and ontogenic roots, yet have developed reciprocally. His stance is the opposite of Piaget's, who believed that both capacities, mental abstraction and language, were directly related. Vygotsky also critiqued behaviorist philosophies of language, finding that they highlighted types of conduct which were conditioned. For Vygotsky, the functions of thinking and language develop independently. He did acknowledge a close relationship between both, based on developmental mental stages, and argued that language could determine the development of thought and other cognitive capabilities, such as abstract knowledge.

Many contemporary experts agree not only that there are innate predispositions toward developing language in humans but also that some language competencies are independent from general cognition. Today, a variety of disciplines, from philosophy to the neurosciences, agree that language is a combination of many different abilities and processes. Neuroscience research, for example, posits that most of these processes may be studied independently of each other, which takes after Vygotsky's proposals; however, this relative autonomy is also based on understanding language as a composite of structures and subprocesses that enable the production of language, knowledge, and related cognitive processes.

Abstract Knowledge and Moral Sense

Many thinkers have long been concerned with the relationship between abstract knowledge and the development of a moral or ethical sense. Although specific differences between morals and ethics exist,

both are concerned with principles, rules of conduct, or habits with respect to correct behavior and with the consequences of wrongful actions. Some experts have taken a biological approach. Recently, some neuroscientists located the seat of concrete thinking and morality—understood as rule-following—in the frontal lobes or prefrontal area of the brain. Others, from a psychological standpoint, examine how people's level of abstract thinking impacts how they develop morality and an ethical sense, with both understood by experts as a kind of abstract knowledge. Some explain that, with higher abstract skills, people become better at looking at the big picture and visualizing myriad and far-reaching consequences. Moreover, people better skilled at abstract thinking are more prone to considering fairness and harm as the basis of ethical norms. Recently, researchers found that when people are thinking abstractly, they are likelier to make judgments on the basis of core values that are applied repeatedly and across many contexts. A recent Yale University psychological study found that participants aligned with the idea that concern about justice and welfare are the core of moral values, particularly when engaging in higher level abstract thinking, such as considering far-reaching consequences or the *big picture* perspective. Views about harm and fairness are elements of human morality that prove more enduring across time than obedience to authority and in-group loyalties, which rely on concrete knowledge. Authority and in-group dynamics are much more contextual factors; that is, based on temporal or situational events, which makes them linked more closely to concrete thinking.

Moral foundations, however, transcend basic concepts of harm and fairness and physical brain functions and capabilities. The issue of moral knowledge and its relationship to abstract thinking has been examined from different perspectives by philosophers since the ancient Greeks. More recently, modern philosophers from Immanuel Kant (1724–1804) to Jürgen Habermas have written about abstract knowledge and critical thinking as related to concepts of the common good. Kant advanced the notion of deep abstract thinking in order to develop a personal and political ethics. Critics of this view, however, have argued that a highly abstracted view of morals ignores the concrete circumstances of events and may fall into rigid adherence to abstract principles that do not

take into account nuanced social contexts and complexities. Nevertheless, Kant's work was invaluable in positing that only freely chosen actions can be moral. Humans, however, cannot freely choose any type of conduct that is dictated by laws that they cannot comprehend nor control, which contradicts the biology-based behaviorist model. Because each moral human behavior must be reasoned, in the Kantian view, it places the onus of a moral sense on abstract thinking and knowledge.

Habermas (1929–), like Kant, views morality within a framework of rational principles. Reasoning, one of the abstract modes of thinking, is of essential importance in preparing for moral action. Morality itself is the consequence of acting in a way that reflects a stable will in the midst of constantly changing events. In other words, moral action is based on reasoning and remains constant through different concrete contexts. Other modern thinkers equate ethics with mathematics, in that both are based on abstract concepts and universal principles. It is important to note, however, that applied ethics are always related to real cases, which brings individuals back to the concrete and contextual.

Most experts believe that in order to develop an ethical and socially responsible citizenry, it is necessary to cultivate not only their critical thinking skills, foresight capabilities, and moral imaginations but also the tools to act on what is ethically correct. There are concrete ways of developing these skills, such as engaging in punishment and rewards in order to foster mechanical learning of what is right and wrong. These, however, can go only so far, and often fail to provide individuals with the information, knowledge, and abstract thinking skills necessary for more profound moral decisions. It is also important to cultivate the abstract knowledge, concepts of justice and democracy, for example, that fosters the desire to be an ethical citizen of the world. Critical thinking skills, which are fundamentally based on abstract capacity, are crucial to developing an ethical education, moral will, and imagination or abstract knowledge necessary to engage in empathetic interaction with others.

Although most scholars often separate the fundamental thinking processes into concrete and abstract, they are not strictly separated. People often begin a thought process by engaging in concrete thinking skills and need concrete examples in order to

understand abstract concepts. Terms such as freedom, morality, racism, hate, and others are highly abstract and can mean different things in different social contexts. It is necessary to contextualize with concrete examples so that people may understand how the concept is being used. The human mind is formed by a network of very rich and complex processes; striving to separate clearly its myriad thought forms and activities is often nearly impossible.

Developing Abstract Knowledge

Thinking strategies, from abstraction and generalization to synthesis and evaluation, serve humans in interpreting reality. Higher levels of thinking depend upon skillfully combining concrete and abstract mental processes; that is, transferring what is learned in one context to another. For example, learning the categorization of elements in a science class or the organization of an essay in an English class may be skills used in organizing information and content in a social science class. Some individuals are more skillful than others at building abstract knowledge, which may be the result of a wide range of factors. Some of them may be to natural predispositions, and extreme cases of difficulty in abstract thinking may be related to mental illness, frontal lobe injury, and other impairments. While there is no known specific treatment that will develop all forms of abstract thinking, some experts posit that it is possible to practice some exercises that may help, such as logic problems, mathematics puzzles, and other brain teasers. However, these are of limited use. The brain and mind processes are highly complex, and while an individual may be a proficient abstract thinker in one area, they may remain a very concrete thinker in others. Thus, somebody who excels at mathematics may have difficulty in literary theory. Exercises that help in one area of abstract thinking may not necessarily improve in others. Other strategies that may help with improving abstract thinking skills are memory problems, problem-solving, organization techniques, and cognition.

In some cases, people can adjust the environment to help concrete thinkers better understand information and content. Some of these include adopting concrete language, avoiding metaphorical concepts or deep levels of abstraction, and always explaining abstract concepts in relation to concrete

examples. For example, a complex social hierarchy system in a different part of the world, which may seem extremely complex to a concrete thinker, can be made easier to comprehend by explaining it in terms of kin or social systems with which they are familiar. It is important to understand, however, that for a very concrete thinker, the progress achieved by these exercises may not translate to other content, academic, or social areas. It does, however, help the individual build up their repository of abstract knowledge in at least one area.

It is a common assumption, to a certain extent, that before the inception of adolescence, very young children do not have any ability to engage in abstract thinking. This is a misconception. Young children begin to develop abstract thinking skills and knowledge through their early pre-school years. One of the ways in which they develop these is through pretend play, which can be taken to a high level of abstraction. For example, children may take an ordinary object to represent something entirely different. They can also engage in complex role-playing conducts which are, basically, variants of symbolic actions. Children learn how to read and to add and subtract by way of symbols, developing the skills that will help them later on to develop more abstract reading comprehension and mathematical skills, such as problem-solving, algebra, and calculus. There are many ways in which children engage in abstract knowledge building. Alphabets and numbers are, after all, symbols. They represent something other than their form, such as sounds and quantities. Children also draw doodles they take to represent something else. When told stories, they can imagine scenes and offer different takes on the storyline. All of these skills involve abstract thinking, which are the necessary stepping-stones to developing strong problem-solving and critical thinking skills later on. It will also be important for developing empathy and a moral self. Abstract skills enable individuals to place themselves in the position of other people and imagine their suffering; in other words, they help develop empathy. These are some of the reasons why experts recommend encouraging and aiding young children to develop abstract thinking skills through a variety of educational programs, games, and exercises.

Trudy Mercadal

See also Abstraction; Conceptual Analysis; Evolutionary Psychology; Knowledge, Logic, Formal and Informal; Thought Experiments, Scientific and Philosophical

Further Readings

- Bickerton, D. (2014). *More than nature needs: Language, mind, evolution*. Harvard University Press.
- Bjorklund, D. F. (2011). *Children's thinking*. Cengage Learning.
- Brockman, J. (2013). *Thinking: The new science of decision-making, problem-solving and prediction*. Harper Perennial.
- Chafee, J. (2012). *The philosopher's way: Thinking critically about profound ideas*. Pearson.
- Cottrell, S. (2012). *Critical thinking skills: Developing effective analysis and argument*. Palgrave Macmillan.
- Drapeau, P. (2014). *Sparking student creativity: Practical ways to promote innovative thinking and problem solving*. Association for Supervision & Curriculum Development.
- Greene, J. (2013). *Moral tribes: Emotion, reason, and the gap between us and them*. Penguin.
- Kenrick, D. T., & Griskevicius, V. (2013). *The rational animal: How evolution made us smarter than we think*. Basic Books.
- Martin, M. W. (2012). *Cognition*. Wiley.
- Napier, J. L., & Luguri, J. B. (2013, November). Moral mind-sets: Abstract thinking increases preference for "Individualizing" over "Binding" moral foundations. *Social Psychological and Personality Science*, 4, 754–759.
- Tomasello, M. (2014). *A natural history of human thinking*. Harvard University Press.
- Yivisaker, M., Hibbard, M., & Feeney, T. (2006). What are concrete and abstract thinking? *LearNet*. Retrieved October 28, 2019, from http://www.projectlearn.net.org/tutorials/concrete_vs_abstract_thinking.html

ABSTRACTION

Abstraction is defined as the development of cognitive signifiers that represent physical objects, actions, or constructs. The word *chair*, for example, might conjure up an image in the mind of a real-world object, a chair. The mind's chair is an abstraction while the physical chair is not, they are different in that one is an approximation, a set

of rules, and the other is an entity. Abstraction is the development of a prototype, or set of rules, that may focus only on a limited set of details. Therefore, it is also integral to the understanding of any theory. We cannot discuss evolution, morality, or quantum dynamics without abstraction. Abstraction is not, however, just a mental process, it can be physical as it is in language, math, computer science, and genetics. Abstraction is sometimes defined as mapping a representation of a problem within another representation. Abstractions are meant to provide insights and clarify a series of observation or to provide a framework for future data analysis. As observations morph, based on an increasing number of documented exceptions, we modify the model little by little until eventually a new model is devised which supplants or augments the original model. This new prototype becomes an added tool with which to measure reality. This entry focuses on defining abstraction through a series of historical usages, defines several schools of thought regarding abstraction, and addresses the concept of abstraction in several commonly studied theoretical fields.

The Theory of Abstraction

It is a little circular that we must have a theory regarding abstraction, as abstraction is a process by which our minds develop theories, that is, by the organization and categorizing of subjects. The term itself is derived from the Latin word *abstrahere* or *to draw away from*. Within the definition of abstraction, we must presume that anything that we look at and interpret with our mind is an abstraction and therefore, so too is the theory to define the term *abstraction*. Abstraction is, in simplest terms, to distill what is generalizable and important about a subject. Our mind does this all the time, in processes that we have only begun to define scientifically. To clarify this concept, let us go back as early as we can in human history. Paintings on cave walls may be one of our first look at human abstraction: a man, a killing tool, and an animal painted on a cave wall. What we have presented is a prototype of the simplest elements necessary for *the hunt*, an abstraction. The three objects—man, tool, and animal as target—existed first, and from it we abstracted the hunt. It

is a distillation of something general, with superfluous details taken out. We can see this in the picture, and we can readily understand it, but it does not show us reality, it shows us an approximation of reality which we call abstraction. Terminology is important when we describe the subjects of abstractions; an abstraction is said to be *conceptual* if it abstracts a concept and *objectual* if it abstracts an object. Therefore, we know that abstraction can indicate at least two types of subjects: objects and concepts. But more than just a generalization of a subject, abstraction is an attempt to define the essential in a subject, winnowing it down to only what is necessary. The following sections aim to describe each of these principles, generalization and essentials, through a survey of historical perspectives.

Language

Language is the most primal and fundamental example of symbolic abstraction. Human languages are at the root of all other complex abstract modeling. Words are the representations of subjects; concepts and objects as encountered above. They reside both in our minds and on our tongues. Linguists believe that abstraction is only possible as a consequence of acquired language, that is, verbalizations. The concept of *red* is an abstraction, not possible, for instance, without verbal ability in which all objects that are red can be divided from other objects only by this shared feature. Before language, one can argue that there were apples, but not *red* apples or *green* apples, as the feature red could not be segregated from the object by a mental representation until it could be generalized to define a specific detail. Written language in turn made visual symbols of sounds, further categorizing the detailed parts of words. Natural language constructs used in language and linguistics are a well-established and formalized form of symbolic concept and object abstraction. Work from the visual languages has suggested that abstraction occurs within this framework as well. A word or hand gesture is used as a representation of a concept. Furthermore, in linguistics, the separation of concepts into deeper categorization, described as the parts of speech, allows for even more elegant rules to be derived about the nature of language itself as well as to foster easier

learning of syntax and grammar. Generally, knowledge acquisition via model construction is accomplished through analysis of verbal or written terminology; in other words, the abstraction of language, as we will see from the discussion of the historical root of abstraction described in the next section.

Ancient Abstractions

The ancient Greek philosopher Parmenides of Elea, in his one surviving work, focused on the power and human tendency toward abstraction. In the poem he talks of the *way of truth* and the *way of opinion*. Truth, he says, is timeless, unchanging, and *what-is*, but opinion, he says, is the way our sensory faculties lead to constructs, mostly false. Parmenides's *opinion* is what we call abstraction, and like his view on *opinion*, what we construct though abstraction is mostly false, or at least approximate. Later we will see how nature and then science have played a role in rectifying at least some of the falsehoods in our abstract models. The Athenians attempted to make a world that was as *real* as possible. Paul Feyerabend described the Athenian view of abstraction in five principles from Aristotle. The first is that the boundaries between reality and appearance cannot be established scientifically. There is a normative or *existential* component. Second, the normative component leads to debate as to the true nature of reality. Third, different lifestyles lead to different interpretations of reality. Fourth, science contains different traditions, and scientific viewpoints are not the only viewpoint. Indeed, many nonscientific cultures suggest that there is more than one way to define a reality that works. And finally, the fifth premise is that science is incomplete and fragmented. What this means is that to the ancient Athenians the world is highly parsed, and no unified view of reality exists. Today we just have more elaborate fragments. Reality then, to the ancients, is an abstraction seen differently depending on what details we make important.

Of note, it was the theological scholar William of Ockham, known for the principle referred to as *Occam's razor* (viz., explanations that posit fewer entities are to be preferred to explanations that posit more), who pioneered the field of *nominalism*, in which he claims that universals, generalizations, are the products of abstraction and have no

extra-mental existence. Nominalism is a position that suggests that only specific objects exist but not universal entities; this stands in contrast to the position of *nihilism*, which denies that even simple objects exist, only mental images do. Finally, *realism* is in contrast to both nominalism and nihilism, in that it suggests that all things are physical, not mental. In these metaphysical world views, we are confronted with the notion of a mental and physical plane. Each philosophy, however, intimates the existence of the abstract, with varying degrees of grounding in the features of the physical world. Though greatly simplified in this discussion, abstractions are considered the properties of real objects in realism, mental interpretations in nominalism, and the only thing that exists in nihilism. Abstraction in this sense means that objects and concepts may not exist before we sense them or at least as we sense them. The idea of something from nothing has been a fundamental conflict between secular and religious scholars, but with the concept of abstraction we find the ability to take not only what is real and formulate nonexistent prototypes (i.e., universals), but also the ability to create patterns, where before there was nothing at all.

Natural and Scientific Abstraction

The behaviorist view is that abstraction only has bearing on reality in so far as it affects actions. Behaviors can be predicated by abstract thought, if we know the abstract model utilized, but it is the outcome that matters. Unlike the aforementioned philosophical viewpoints, *naturalists* believe that nothing exists outside of space-time. From this perspective, all abstractions exist within and as a part of natural phenomena; in other words, everything can be reduced to a physical account, even conceptions. Other viewpoints, such as Ockham's, which state that there are certain things that cannot be rationally examined, are not interpretable by our rule sticks and therefore not useful in theory design. Nature, however, played an important role in the development of abstract worldviews before a systematic scientific perspective came about. As mentioned previously, Aristotle believed that while many viewpoints existed, they were all required to be successful if they were to persist. Abstractions were valuable in so far as they helped to predict or guide successful human behavior. In

many cases, these early abstractions were the models for morality and ethics on the cultural level. For example, the perception of the Egyptians that beer was a product of the gods led to widespread consumption, which in turn kept bacteria at bay and paid for the building of the pyramids. Abstraction, from this perspective, is a view of the world that allows for better control of it.

Another natural view of abstraction comes from the art world, where abstraction has been emphasized as an inspiration for all artistic forms. Abstract art is not, as the uninformed may assume, a mass of scribbles on a canvas but is inspired by an artistic attempt to eliminate unnecessary elements while retaining the universal. In art theory, abstraction is taught to aspiring artists as the process of stripping down to the essential. It is thought that some abstraction occurs in every form of art, and it is this essence that the artist is intending to elucidate with their work. This is in effect a reversal of nature impinging on our internal model; in art we have an external model reflecting nature's essence. In both domains, we find the first seeds of scientific abstraction, that is, the beginning of detailed perspectives in the development of better modeling systems.

Science is the way we try to decipher our world into manageable chunks. The formalization of abstract thought into symbolic chunks called *language* paved the way for a deeper analysis of reality. We have heard that Aristotle believed that science could not fully interpret reality, that Ockham believed that there were ideas that could not be divined through rational thought, but despite this, the environment influences our model and artists have tried to distill the fundamental in their works. Likewise, scientists have continued to push the bevel incrementally forward. From a strictly scientific perspective, abstractions are the subdivision of data packets, information, into their smallest unit of representation. By systematically categorizing, science has devised abstractions of our world that are superior to the abstractions of the past. Scientific abstraction is meant to better predict the outcomes of our interaction with reality; again, this is achieved by narrowing down to what is essential or universal in a process. One way to do this was to attempt to remove ourselves from the equation, to form a detachment from the principles with which we wish to form theories. The scientific method was one way to achieve

that, and this has been done at least twice more in human history with great success. The first was with human language, described earlier in this entry, and the second was with mathematics. Both of these symbolic systems, which distill objects and concepts into universal icons, are fundamentally systems of abstraction.

Mathematics

Though some have argued that mathematics is a self-contained system, no system that is used to represent another system can be defined as anything other than an abstraction. Mathematics is the second most pivotal example of abstract modeling devised by man and has allowed us to visualize, though symbol, scales of magnitude far greater and far smaller than our bodies could comprehend otherwise. From it, we have developed rocket propulsion, physics, and computer science. While arguably more or less important than language in understanding reality, math most likely could not have been derived without it. Indeed, mathematics is perhaps a better abstract modeling tool than language because it is more detached from human emotions and anthropomorphic properties.

Mathematical abstraction requires utilization of everyday words, but in a very precisely defined relationship to mathematically unique symbols. Additionally, mathematics consists of sets of rules for operating on mathematical objects. Therefore, symbolic objects and concepts are manipulated by a specific set of rules to form a self-contained abstract modeling system. Like all abstract systems, the rules and the subjects are generalized; universals represent other structures. Numbers, for example, were initially mathematical symbols that represented quantity. Later, axiomatic systems arrived that were considered independent of quantity, which emphasized the abstract nature of the mathematical objects and suggested that they may exist apart from external reference. Again, if you recall, abstractions can be representations of real-world objects, such as a *dozen*, or they can be representations of things that do not physically exist; mathematics reminds us of the nihilistic view that nothing physically exists.

One interesting mathematical concept, which we will see again in computer science, is the mathematical *map*. A mathematical map is a function used to describe a class by a specific parameter, such as the

linear transformation in algebra. As such, a mathematical map is an abstraction, which is defined by mathematical abstractions. There are of course many other cases of recursive abstraction. It is math's real-world utility that makes it a powerful abstraction tool, from calculating a monthly budget without cash in front of you to developing a video game using a simulated physics engine.

Psychology

Schema-abstraction, defined in the psychology literature, is the process of learning concepts from examples. Schema-abstraction theories assume that some information is abstracted during learning and stored for classification of new examples. And these theories are distinguished from one another based on how information is characterized, retained, and utilized in the classification of new incidences. In schema-abstraction studies, either a *prototype* model, in which an abstract prototype is constructed and all other examples are compared to it, or an *instance-only* model, in which no abstraction is performed and only training examples are stored, is postulated. Determining the difference between these types of interpretive models has been conducted within psychology research in collaboration with computer science, where the research tools are developed. For example, using a computer mapping program, a series of experiments meant to examine what effect sequence presentation has on conceptual learning was conducted, whereby subjects are asked to devise a rule based on observation of a series of figures given in a specific sequence. However, an explicit definition of the abstraction process has made these studies difficult. For instance, some have argued that learning is dependent on the sequence of presentation of events; others, however, have only seen weak results in experiments looking to replicate this principle. The primary function of schema-abstraction theory investigation is to attempt to delineate which aspects of a model are generalized and when this generalization takes place.

Computer Science

The use of computer programming to generate a number of plausible mappings (*gmaps*) of symbolic object associations has been used in the study

of the cognitive abstraction processes described above. Mapping, in computer science, is derived from the mathematics and is used to describe the process of converting associative arrays of data from one form to another. Typically, these programs consist of several modules including the map generator and the structural evaluator, which acts to providing alternate gmap evaluation scores as well as other descriptive modeling tools. Studies in this area have loosely defined abstraction strategies into two classes *limiting case* models, including exemplars only, and *combination* models, including an IF-THEN-Rule Generator. In computer-aided modeling, investigators can investigate several models at the same time. Each model predicts a different type of mental rule formation and can therefore help to define what part of an observation is being processed as an abstraction. Developing investigation tools is only part of interest for computer scientists in abstraction modeling. The binary code of 1s and 0s represents an attempt to distill the information found in analog data into a simpler form. Combining mathematical rules and computer science has led to the creation of intricate simulated environments, sometimes called *engines*, which have been used to analyze everything from engineering to biology. Computer scientists have used abstractions from web browsing histories to predict customer activity, which has been used by search engine companies to target products to consumers. Another lucrative area of research in abstraction is in the field of artificial intelligence (AI). Indeed, understanding the pieces of information used in conceptual abstraction can influence the design of computer systems to create more efficient information processing, one of many goals in creating advanced AI.

From Cognition to AI

Simpler models act as the starting point for more complex modeling. But in contrast to what one might expect, the derivative symbolism from the simpler models allows us to formulate complex models in simpler terms. Like the mathematical map, this is the nature of abstraction. One abstraction can become embedded in another abstraction, reducing the size of the data. By ignoring the irrelevant details, abstract models take many subjects and combine them into one subject. We have seen

this in philosophy, where Ockham suggested that the simplest solution usually fits best and called that solution *God*. We have seen this in language, whereby one word, such as *chair*, defines a whole category of objects. We have seen it in math, where one number, for example 5, is the defining factor that generalizes a multitude of instances. We have seen how it can be applied to schema-abstraction models in which either a prototype or exemplar exists. We know that computer scientists are using abstractions today to track the behavior of individuals and market products to each individual accordingly. But these are just the beginning. Abstract models beget more abstract models, which again work to condense multiple objects into symbols representative of them all. The biological process of cognition is thought to use abstraction, which is why it created language and math, and is rooted in the symbolic DNA code of four base pairs, which represent 20 amino acids, from which countless protein structures are generated. Several structures of the human brain are called *pattern generators*. Humans intuitively try to associate objects and concepts into clusters called *patterns*; in other words, we naturally abstract. Perhaps now we can see the purpose of abstraction, a storable code that represents a much more complex system. Data are easier to manage in this simpler form and take up less space. Indeed, humans have extrapolated that the ultimate abstractions would be the ones that define all subjects as one essential category: the unified theory, the Big Bang, God, and so forth. Finally, at the forefront of science, individuals are piecing together the processes found in numerous fields to generate an abstraction of human thought that can fuel an artificial mind. Whereas human cognition is still hundreds of orders of magnitude more efficient than the currently most advanced computer operating system, the growing field of AI has an overwhelming need to define abstractions to fit more information into a finite system. These are the yet-to-be-defined abstract information processing algorithms.

Conclusion

Abstraction, as explained in the preceding sections of this entry, is the process of condensing relevant information and disregarding irrelevant information. It is used in commonsense reasoning

in order to simplify conceptualizations and to focus attention on specific details. Our processing of reality occurs in the form of abstraction. Furthermore, language and mathematics, two formalized abstract models, have led in turn to refinement of our abstract characterizations and were required for advanced scientific discoveries in biology, psychology, and computer science. To abstract is to focus on establishing the essential in every system. This is the nature of abstraction: to ever refine and reduce subjects to their simplest and most essential characterizations. Abstraction therefore is one of the most useful tools in theoretical analysis.

Ian C. Clift

See also Artificial Intelligence; Information Theory; Mental Models; Metaphysics; Data Models

Further Readings

- Feyerabend, P. (1999). *Conquest of abundance: A tale of abstraction versus the richness of being*. University of Chicago Press.
- Fine, K. (2002). *The limits of abstraction*. Oxford University Press.
- Loux, M. J. (Ed.). (1979). *The possible and the actual: Readings in the metaphysics of modality*. Cornell University Press.
- Wielinga, B., Boose, J., Gaines, B., Schreiber, G., & Van Someren, M. (Eds.). (1990). *Current trends in knowledge acquisition: Volume 8 of Frontiers in artificial intelligence and applications*. IOS Press.
- Zucker, J.-D. (2003). A grounded theory of abstraction in artificial intelligence. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, 358(1435), 1293–1309. <https://doi.org/10.1098/rstb.2003.1308>

ACCURACY

The concept of accuracy is central to the scientific method; this article focuses on how it is employed in scientific practice. In common parlance, a statement is said to be accurate when it is factual. In turn, a statement is factual when it is expressing a proposition that is in agreement with facts or actual states of affairs. In that sense, asserting that

a statement is accurate is equivalent to asserting that the statement is true. Moreover, the term *accuracy* is often used as a quasi-synonym for other terms such as exact, precise, correct, literal, and similar expressions in the same semantic field. However, in scientific parlance, there are significant differences between the meanings of those terms.

The concept of accuracy is closest to what philosophers of science call *approximate truth*. (Other terms used in philosophy of science with similar meanings include truthlikeness and verisimilarity.) This is reflected in the normative terminological proposals of the International Organization for Standardization (ISO, 2012) where the term *trueness* is used to describe the accuracy of a measurement method ($\cdot$). (The ISO standards also uses the term *precision* to describe accuracy. However, we will not follow this usage here since it conflicts with standard usage in mathematics and computer science, where *accuracy* and *precision* are sharply distinguished.) As such, accuracy is a concept that is related to other methodological concepts such as precision, error, and uncertainty, and this article will sketch how the relation is fleshed out in experimental and computational contexts, which are the two main scientific contexts in which assessments of accuracy are required.

Accuracy is a concept used to characterize statements *semantically*, just like approximate truth. Thus, whereas truth (and accuracy in its colloquial sense) is used to provide a *binary* semantic evaluation—true or false, with no admissible intermediate—accuracy and approximate truth allow for intermediary semantic assessments that admit of degrees. But due to its intrinsic vagueness and lack of general criteria of application, the concept of approximate truth has received some harsh criticisms in the philosophy of science literature. For instance, Laudan (1981) has argued that discussions relying on approximate truth are “just so much mumbo-jumbo.” Indeed, what would it mean for the statement *Orgone exists* to be approximately true? Such qualms, however, are dissipated once we observe that approximate truth and accuracy are concepts that do not apply to the semantic evaluation of all kinds of statements.

The scope of application of the concept of accuracy can be rigorously delineated using Carnap’s distinction between classificatory (also

known as *qualitative*), comparative, and quantitative concepts. Firstly, a classificatory concept is a concept that places an object within a certain class. For instance, when we assert that whales are mammals or that electrons are fermions, or when we make an existentially quantified assertion such as “*dark matter exists*,” we are making simple claims about classificatory concepts. Such concepts can typically only be meaningfully evaluated semantically using a binary scale—true or false. Thus, neither the concept of accuracy nor that of approximate truth apply. Secondly, a comparative concept is one for which it is meaningful to provide an ordinal ranking of the degree to which the concept applies. For instance, if we consider the classificatory concept *tall*, we could only assert that a given person is tall or not. With the comparative concept *taller than*, we are now in a position to say that individual *a* is taller than *b*, that *b* is taller than *c*, and so on. However, quantitative concepts are purely ordinal. Thus, the concept *taller than* does not offer resources to assess how much taller *a* is, *b* is, and so on.

Finally, quantitative concepts are those for which there is a numerical value on some scale that is used to determine whether the concept applies. If we consider the concept of distance, we would mathematically specify with a number or a function what the distance of a line segment is. There are many subtleties and complications involved with the selection of a proper scale to assign numerical values as Scott and Suppes make clear in their 1958 essay “Foundational aspects of theories of measurement”. However, the point that is important for our discussion is that if we assert

The line segment has length M (1)

when in fact it would be true to say that

The line segment has length L , (2)

we can give a semantic assessment of statement (1) more refined than just observing that it is false. To the extent that M is close to L , we will say that statement (1) is more or less accurate. We then define the quantity Δ

$$\Delta = |M - L| \quad (3)$$

as the *error*. The smaller the error, the more accurate a statement is. Thus, statements containing quantitative concepts are within the scope of

application of the semantic concepts of accuracy and approximate truth.

However, there is a difficulty related to the assessment of the accuracy of a statement that fundamentally entangles it with a related epistemological problem. Indeed, if we knew what the exact (or true) value of the quantitative parameter in the first place, we would have asserted the true statement. Thus, the problem is that we need to assess the accuracy of statements without knowing which value is the exact one. This epistemological problem is central to two methodologically important branches of science, namely, scientific computing and experimental measurement. We discuss them in the remaining two sections of this article.

Accuracy of Computation

The concept of accuracy plays an essential role in computational mathematics because exact solutions to the mathematical equations arising from the construction of realistic models typically do not afford exact solutions. Numerical algorithms are then used to extract information about what the solution to model equations might look like. However, since the computed solutions are typically not exact, an assessment of accuracy is required to validate results.

Computational errors are essentially of three types. Truncation error amounts to replacing functions $f(x)$ (often characterizing vector fields) and integrals $\int f(x)dx$ (often characterizing the motion of a body in phase space) by truncated asymptotic series in a perturbation parameter ε , that is,

$$f(x, \varepsilon) = \sum_{k=0}^N f_k(x) \phi_k(\varepsilon),$$

for some collection of gauge functions $\{\phi_i\}_{0 \leq i \leq N}$ (see the entry on Perturbation Theory in this encyclopedia). Expressions of this sort have to be truncated, since we often have no closed-form solutions, and it is impossible to add an infinite number of terms in series. Secondly, discretization error is the error incurred by replacing a continuous parametrized flow $\dot{x} = f(t, x(t); \mu)$ in phase space by a discrete map of the form

$$x_{k+1} = \Phi(t_k, x_k, \dots, x_0, h, f, \mu).$$

This substitution is the basis for most methods of numerical differentiation and integration. Finally, we typically do not compute the value of functions using field arithmetic (e.g., the familiar arithmetic of real numbers), since computers cannot handle such infinitary entities. Thus, it is replaced with a finite computer arithmetic known as *floating-point arithmetic* (see Figure 1).

In essence, it involves replacing the real line by a *floating-point number line*. As we see in Figure 1(b), it is not really a line; this is why it is important to consider the role of roundoff error in mathematical representation. All of these computational approximations are made because we can only execute finite, discrete operations on a digital computer.

There are three ways of measuring computational error. (For a more extensive explication of these notions, see Corless and Fillion (2013), Higham (2002), or Deuffhard and Hohmann (2003).) Mathematical problems can be thought of as maps from a set of input data to a set of

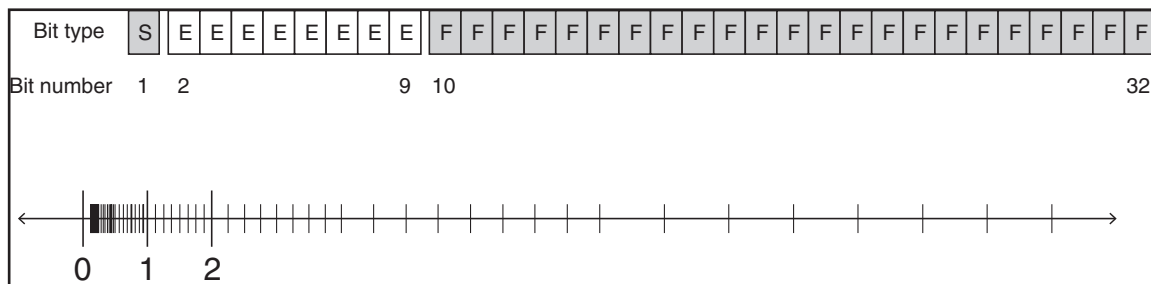


Figure 1 Floating-Point Numbers. (a) Structure of Single-Precision Floating-Point Numbers. There Is a Bit for the Sign, 8 Bits for the Exponent, and the Remaining 23 Bits Are for the Fractional Part. (b) Discreteness of the Floating-Point Number Line, With Variable Density

output data, that is, as $\varphi : \mathcal{I} \rightarrow \mathcal{O}$. The mapping itself will typically be given as

$$x \xrightarrow{\varphi} \{y \mid \phi(x, y) = 0\}, \quad (4)$$

where x is the input data, y is the solution sought, φ is the defining function, and $\phi(x, y) = 0$ is the defining equation (or condition). However, as we have seen, when we use a computer for computations, we engineer a modified problem $\hat{\phi}$ so that it can be implemented and efficiently executed on a digital computer using schemata for discretization, truncated, and rounding off. In general, for a problem such as the one in equation (4), the difference between the exact solution y and the approximate solution $\hat{y}$ will be denoted Δy and will be called the *forward error*. The concept of forward error is typically the one that is referred to when people talk about error; when they refer to a computed solution as being *accurate* or *approximately true*, one means that the forward error is small.

Even if the forward error is the most common, in many contexts this is by no means the most insightful, because we have no direct way of calculating Δ when we don't know what y is. For this reason, there are two other measures of error known as the *backward error* and the *residual*. (For the former, see Fillion and Corless (2014), and for the latter, see the article Perturbation Theory in this encyclopedia.) However, whatever measure of error is used, assessment of accuracy is typically done by reference to the forward error. Since it is typically impossible to obtain an exact quantification of the error, numerical analysts typically analyze the algorithms used and prove that the error must be smaller than a certain quantity or function, known as an *error bound*. As a result, even if the exact error is not known, we know that it lies within a certain interval. If the interval is sufficiently restrictive for the intended purposes, then the computed solution is judged sufficiently accurate.

Accuracy of Measurements

The construction of a model susceptible of generating prediction (or retrodiction) and explanation of the behavior of a system requires the

specification of the value for some parameters. Those parameters may be the primitive quantities in which the states of the system are expressed—examples include position, momentum, temperature, pressure, spin, and so on—or some other derived quantity. In either case, the value of this parameter will be supplied on the basis of *measurements*. Nonetheless, no scientist would argue with the fact that all experimental data have some degree of imperfection, in that experimental results always contain errors. As a result, numerical values gathered in experiments are always likely to be wrong (i.e., inexact). This, however, does not imply that the values reported are bad, for they may convey entirely satisfactory information if the measurements are sufficiently accurate. To ensure that inexact values reported are informative, scientists have to diagnose the possible sources of measurement error and must try to design an experimental setup that will ensure that the error is minimized (or satisfactorily small, given what is already known). That is, one must establish the value of a parameter with a methodology that will also make it possible to assess the accuracy of the value.

However, here again it is not possible to exactly quantify the measurement error directly for the value of the parameter being measured is unknown. The key to assessing the accuracy is not to directly quantify the error but to specify an interval within which the error is. Thus, the role of measurements of parameters is to determine (1) a value of the parameter and (2) an estimate of the uncertainty associated with the measurement. The central concepts involved in a theory of measurement are thus the concepts of accuracy, error, and uncertainty.

The discussion to follow is based on the so-called *GUM approach* to the theory of measurement (“GUM” stands for “Guide to the expression of Uncertainty in Measurement”). The relevant technical documentation includes the guidelines of the National Institute for Standards and Technology and the technical reports from the Bureau International des Poids et Mesures. In what follows, the terminology is used according to the International Organization for Standardization (2004).

A measurement is a process involving a *system* and an *apparatus*. The quantity that is subject to

measurement is called the *measurand*. The value resulting from the measurement using a certain apparatus is known as the *indication value*. The difference between the value of the measurand (what is often called the *true value*) and the indication value is the *error*. Again, error is a *semantic* notion, relating to a *matter of fact* relating two numerical values. It is not about what we know, ignore, wish to know, or even can know. It must be distinguished from the *epistemological* notion of uncertainty, since “the result of a measurement after correction can unknowably be very close to the unknown value of the measurand, and thus have negligible error, even though it might have a large uncertainty.” It is important to stress the distinction since they may *appear* identical as both are understood as intervals within which the *true value* lies. All the scientist can do is to provide an *estimate* of the error based on what is known about the system and the measurement apparatus. (“In general, the error of measurement is unknown because the value of the measurand is unknown. However, the uncertainty of the result of a measurement may be evaluated.”)

This is why, in modern expositions of the theory of measurement, the role of uncertainty is given priority over that of error for the formulation of methodological rules. The first step toward a correct estimation of the uncertainty of the results of a measurement is a diagnosis of the possible sources of measurement error, which are of two kinds: (1) random error and (2) systematic error.

They are also often referred to as *Type A* and *Type B* error, respectively. On the one hand, *random errors* are unpredictable. They are variations in the measurements that the experimenter cannot control (or can control only very limitedly). In terms of probability, it is an error that is just as likely to be above or below the real value; so, for random error, averaging a large number of measured values should, in principle, largely reduce the magnitude of the error. On the other hand, *systematic error* cannot be controlled as random error, that is, averaging will be of no help. Systematic error is caused by the design of the experiment. Its impact can only be alleviated by modifying the design; however, it is

often very hard to find a setup that has no systematic error. See Table 1 for examples. Systematic error is particularly problematic, from an epistemological point of view, because there is no way to determine whether there is a systematic error. (“Like the value of the measurand, systematic error and its causes cannot be completely known.”)

Table 1 Examples of Random and Systematic Error

<i>Random Error</i>	<i>Systematic Error</i>
Vibration in the floor → fluctuation in balance	During the time required to measure the mass of a fluid, some evaporates
Air currents → fluctuation in balance	During the time required to measure length, the temperature is not controlled and changes
Electrical noise in a multimeter	Miscalibrated balance will cause all the measured masses to be wrong

Each of the uncertainty components that contribute to the uncertainty of the measurement are represented by an estimated standard deviation, termed *standard uncertainty*. This standard deviation may be or may not be evaluated statistically, as Case A or B may be. We talk of *Type A* and *Type B* evaluation of uncertainty. Type A uses any appropriate statistical method. However, the procedure is not as straightforward for the assessment of Type B uncertainty components:

A Type B evaluation of standard uncertainty is usually based on scientific judgment using all the relevant information available, which may include

- previous measurement data,
- experience with, or general knowledge of, the behavior and property of relevant materials and instruments,
- manufacturer’s specifications,
- data provided in calibration and other reports, and
- uncertainties assigned to reference data taken from handbooks.

Repeating the experiment can be used to successfully control error arising from random effects. By carefully considering the factors mentioned, *replicating* the experiment can be used to successfully control error arising from systematic effects. (It is important to distinguish repeatability and replicability. Repeatability consists in doing the measurement multiple times with the same apparatus; replicability involves changing the apparatus, each having their own systematic biases.)

The task that remains is to find the *combined standard uncertainty* of a measurement result. The combined standard uncertainty, as its name suggests, combines together all the uncertainty components of Type A and B to generate a total estimate of uncertainty. The usual method used to combine uncertainty is the common statistical method used to combine standard deviations known as *law of propagation of uncertainty*.

We have seen that the concepts of error and uncertainty are intrinsic aspects of the values of the experimentally measured parameters. What does the presence of error and uncertainty imply for the mathematics used to analyze such data? The key challenge here is to develop mathematical techniques that permit us to deal adequately with operations on quantities not known exactly. The most basic approach to do this is that based on *significant figures* (or, alternatively, *significant digits*). The idea is easily understood from the following joke:

Bazooka Joe is showing a friend a fossilized bone. The friend asks how old it is and Bazooka Joe responds that it is one hundred million and three years old. “How do you know that?” asks the friend. Bazooka Joe responds “The museum expert told me it was a hundred million years old and that was three years ago.” (reported by Ruekberg, 1994)

The author then explains the pedagogical relevance of the joke as follows:

The author readily admits that the joke is not *very* funny. That it is funny at all is because even children, just old enough to chew gum without swallowing it, realize that something is wrong about Bazooka Joe’s computation: the accurate three years cannot be added to the ball park figure of a hundred million years. (Ruekberg, 1994)

The computation is wrong even if, arithmetically speaking, it is irreproachable. Thus, when there is uncertainty, there is a methodologically important sense of “correctly using mathematics” that differs from the standard one. Understood as a tool occupying a central place in a strategy for the management of uncertainty, the first role of significant figures is to faithfully report the uncertainty in experimental measurements. More precisely, significant figures are a tool to faithfully report the *accuracy* and the *precision* of the results of a measurement, given the *resolution* of the measuring device. It is important to keep in mind the distinction between accuracy and precision. Accuracy is about having the answer right, that is, about having a small error. Precision is about having many digits. For instance, 3.166666666666667 is a very precise (16 digits) but (for many purposes) very inaccurate (two digits) approximation to π . On the other hand, 4.54×10^{-5} is a not very precise (three digits) but quite accurate (precise to the order 10^{-6}) approximation to e^{-10} . Similarly, a measuring instrument can be very precise, and yet inaccurate. As W. Kahan and Joseph D. Darcy (1998) explain, “[p]recision is to accuracy as intent is to accomplishment.” A basic objective of the use of significant figures is to not be fooled by measurements that are more precise than accurate; thus, a basic rule is to not report results of measurements with more digits than are accurate, for those extra digits would not be significant. Precision is a property of a linguistic object (namely, of the numeral representing a number in a given number system), whereas accuracy is a semantic notion. Limiting the number of figures reported to the significant figures is a way to make the semantics transparent by showing it in the form of the linguistic expressions used to report results.

The resolution of an instrument is the maximum error that the instrument produces under prespecified circumstances (e.g., value range, ambient temperature, humidity, pressure). (The manufacturers of equipment often indicate how accurately and precisely it can measure.) However, the resolution cannot be smaller than the precision of the instrument. To illustrate this point with a simple example, if a ruler’s smallest

division is 1 millimeter, then we cannot specify what a length is by measuring with this ruler is to less than half of a millimeter. What we obtain from an instrument with this precision is a number having the format $x.yyz$ centimeter, where x is the integer part, yy are the certain digits, and z is the uncertain digit (there is only one of those, the last one). The last digit is only an estimation. Moreover, under the prespecified conditions mentioned earlier, if the instrument is properly calibrated, then each digit within the precision of the instrument is taken to be significant. Accordingly, *the significant figures of a number are those digits that carry meaning contributing to its precision*, and indirectly contributing to its accuracy, provided that some assumptions about calibration are satisfied. (Of course, if the prespecified conditions of calibration are not met, then the reasoning does not hold.) Thus, the significant figures of a number are the digits necessary to specify our knowledge of that number's precision; and in nice contexts, this also reveals the accuracy.

Whereas this first role of significant figures is to *represent* uncertainty numerically by imposing conditions on precision, the second role of significant figures is to permit the formulation of computation rules to determine how uncertainty *propagates*. Those rules are typically formulated in terms of limits on the numbers of significant digits (or significant decimal places) that may be retained in order to faithfully track uncertainty. Here are some standard rules:

1. Exact numbers do not affect the number of significant figures.
2. For addition and subtraction, the answer contains the *same number of decimal places* as the least precise operand used in the calculation. The idea is that you cannot add to or subtract from something not known.
3. For multiplication and division, the answer contains the *same number of significant figures* as the least precise operand used in the calculation.
4. For logarithms, only those numbers to the right of the decimal place of the operand count as significant.

Such a set of rules constitute what is called a *significance arithmetic*.

Now, these rules should not be thought of as being perfectly reliable. Rather, they work as rules of thumb. One might be surprised that there are, well into the 20th century, many publications debating how significant digits should be analyzed and understood. This is because for any set of such rules based on significant digits it would be relatively straightforward to generate problematic cases. Thus, we see that as an attempt to provide context-independent syntactic rules meant to support management of uncertainty and its propagation, significance arithmetic has limitations. More refined mathematical methods of assessing error and uncertainty propagation in order to properly assess accuracy in a more robust way include sensitivity analysis and perturbation theory.

Nicolas Fillion

See also Data Models; Measurement; Perturbation Theory

Further Readings

- Carnap, R. (1966). *Philosophical foundations of physics: An introduction to the philosophy of science* (300 pp.). Basic Books.
- Corless, R. M., & Fillion, N. (2013). *A graduate introduction to numerical methods, from the viewpoint of backward error analysis* (868 pp.). Springer.
- Deuffhard, P., & Hohmann, A. (2003). *Numerical analysis in modern scientific computing: an introduction* (Vol. 43). Springer Verlag.
- Fillion, N., & Corless, R. M. (2014). On the epistemological analysis of modeling and computational error in the mathematical sciences. *Synthese*, 191, 1451–1467.
- Higham, N. J. (2002). *Accuracy and stability of numerical algorithms* (2nd ed.). SIAM.
- International Organization for Standardization. (2004). *International vocabulary of basic and general terms in metrology (VIM)*. Technical Report DGUIDE 99999. Revised from the original 1993 edition.
- International Organization for Standardization. (2012). *Accuracy (trueness and precision) of measurement methods and results, Part 1*. Technical Report 5725-1. Revised from the original 1994 edition.
- Joint Committee for Guides in Metrology. (2008). *Evaluation of measurement data—Guide to the*

- expression of uncertainty in measurement*. Technical Report JCGM 100:2008. Bureau International des Poids et Mesures. Revised Edition of GUM 1995. Joint Committee for Guides in Metrology. (2009). *Evaluation of measurement data—an introduction to “guide to the expression of uncertainty in measurement” and related documents*. Technical Report JCGM 104:2009. Bureau International des Poids et Mesures.
- Kahan, W., & Darcy, J. D. (1998). *How Java’s floating-point hurts everyone everywhere* [Presentation]. ACM workshop on Java for high-performance network computing, Stanford University, CA.
- Laudan, L. (1981). A confutation of convergent realism. *Philosophy of Science*, 48, 19–48.
- Ruekberg, B. (1994). A joke based on significant figures. *Journal of Chemical Education*, 71(4), 306.
- Schwartz, L. (1985). Propagation of significant figures. *Journal of Chemical Education*, 62(8), 693.
- Scott, D., & Suppes, P. (1958). Foundational aspects of theories of measurement. *Journal of Symbolic Logic*, 23, 113–128.
- Taylor, B. N., & Kuyatt, C. E. (1994). *Guidelines for evaluating and expressing the uncertainty of nist measurement results*. Technical Report NIST Technical Note 1297, National Institute of Standards and Technology.

ACTOR–NETWORK THEORY

Actor–network theory (ANT) developed within the wider interdisciplinary field of science and technology studies (STS), and more specifically, out of debates within the sociology of scientific knowledge (SSK). SSK, in a series of case studies of scientific controversies, demonstrated that the acquisition of scientific knowledge could not be fully explained by its logic, method, or normative institutions. Taking the work of Thomas Kuhn (1922–1996) as a starting point, SSK demonstrated that the very content of scientific knowledge, including mathematics, was socially constructed. As Bruno Latour (1947–) most explicitly argued, however, SSK inadvertently replaced the natural determinism of the philosophy of science with a social determinism that was no more tenable. In addressing this criticism, ANT’s seminal case studies such as Latour’s *The Pasteurization of France* and Michel Callon’s

“Some Elements of a Sociology of Translation” followed the process of *heterogeneous engineering* by which technoscientific objects and truths were assembled and made durable through networks of human and nonhuman actors. Thus, ANT extended SSK’s symmetry principle, which held that the same types of causes should be used to explain both true and false beliefs, to a *generalized symmetry* that treated humans and nonhumans as actors with agency. Though the exact construal of nonhuman agency within ANT remains subject to debate, largely around whether to preserve intentionality as an analytically distinct feature of human action, ANT immediately gained attention as an approach to studying how *the social* is continuously enacted as relational networks of humans and nonhumans. In ANT’s subsequent articulations as a general framework for studying social order, three of its concepts, *heterogeneity*, *material relationality*, and *punctualization*, initiate a radical reconceptualization of agency, social action, and society itself. This entry addresses each concept, in turn, as well as its consequences for humanistic social theory.

ANT insists that social phenomena are *heterogeneously* constituted. That is, its objects of inquiry—whether scientific facts, nation-states, organizations, and so on—are made up of both human and nonhuman components. ANT emphasizes that a potentially endless array of *actants*, a term borrowed from literary theory and semiotics to designate any human or nonhuman that can accomplish action in a network, may constitute an object of interest. An aspect of ANT’s empirical concern thus lies in identifying, or *following*, the specific actants that are relevant in shaping the phenomena in question. In the case of California state formation, for example, Patrick Carroll argues that it is one of the most productive agricultural centers in the world only so long as levees, dams, irrigation, farmers, metering devices, state agents, and so on, are networked together. Heterogeneity greatly distinguishes ANT from humanist social sciences that, in this example, would overlook the relationship between state territories and their material infrastructures; that is, how the historicity of material culture shapes the state of California. Under an actor–network formulation, then, the state’s ontology is composed of heterogeneous elements rather than viewed as a single,

coherent actor that is distinct from society. So, an initial premise of ANT is that *the social* is made up of a heterogeneous collection of humans and nonhumans, though their relationality is always an empirical question.

To the heterogeneity of the social, add ANT's concern for *material relationality* and this is how ANT proposes to study the social world: not from *a priori*, binary divisions between humans/nonhumans, nature/culture, state/market, and so on, but in terms of how such phenomena are constituted in the continuous unfolding of heterogeneous relations. This second actor–network concept argues that entities have no essential reality apart from the web of relations in which they are produced. In other words, social phenomena are generated by human and nonhuman actor–networks and it is the relationality between heterogeneous elements that ANT's case studies pursue. This concern for relationality orients ANT toward the *way things are*, the ontological character of the heterogeneous actor–networks through which the world is made. ANT argues that what gives objects, forms of knowledge, and other social phenomena their meaning, stability, or power, must be examined in terms of how their *bits and pieces* are held together.

In the decades-old project of damming the Mississippi River, for example, an effective levee-building science develops in relation to the resistance of the river itself. Andrew Pickering captures this relational process as a *dance of agency* between engineers who build dams, a river that resists their control, and the subsequent redesign of their materials. The result is an engineering science that has become increasingly more attuned to the sedimentary composition and currents of the river. ANT case studies emphasize this always contingent, never final, way in which humans and nonhumans form more durable wholes through complex material relations. In this sense, ANT shares an emphasis on the unfolding and processual dimensions of the social world characteristic of poststructural, ethnomethodological, and pragmatist philosophies.

When taken together, heterogeneity and material relationality are significant both for the empirical study of scientific practice as well as for ANT's theory of social action. With respect to the former, ANT suggests that the work of technoscientists is not simply to follow a method, conduct experiments, and reveal the *truth* about the natural

world. Rather, technoscience is a complex and strategic practice in which the diverse materials encountered must be *assembled* into a durable, material network that can overcome unforeseen obstacles. For example, Bruno Latour's seminal case study rejected the familiar story that attributes Louis Pasteur's success to his genius mind. Instead, Latour pointed to Pasteur's technoscientific practice, the material arrangement of a heterogeneous actor–network composed of laboratories, domesticated bacteria, statistics, notebooks, microscopes, and more, as the explanation for his triumph. Pasteur's success, in other words, is an effect of organizing these heterogeneous components into a stable, material arrangement. Thus, in ANT's reformulation of social action, it was not Pasteur's rational mind that distinguished him from other technoscientists, but the collective strength of the heterogeneous actants he *assembled* as a network. This insight further supports the STS thesis that science is a social and cultural activity, and in a final move, further reconceptualizes social action.

The third concept, *punctualization*, refers to the process by which a heterogeneous network stands as a singular and coherent object or actor. It suggests that when *Pasteur-the-great-researcher* is evoked as a scientific hero it is as an effect of his heterogeneous network. Crucially, punctualization entails an oscillation between powerful individuals or objects and the contingent, network-building practices through which they become recognizable as such. The concept can thus be viewed as an alternative to human-centered histories and philosophies of science that tend to *black box* the process of knowledge production; that is, attribute knowledge itself to the brilliance of a solitary individual. As ANT instead directs empirical attention to precisely the activities through which a matter of fact is successfully constructed, it explains Pasteur's success as a heterogeneous and collective achievement, an effect of the associations of multiple heterogeneous actants. Under an actor–network formulation, Pasteur cannot become a brilliant technoscientist of eventual world renown apart from a durable network. Still more radically, Pasteur's achievement *is* the collective achievement of a heterogeneous actor–network for which he, when evoked as a singular individual, becomes a *spokesperson*.

In conclusion, ANT argues that agency, social action, and society itself are empirically complex and in need of explanation. Whereas humanist social sciences tend to privilege a social world made up of intentional, human actors and defined by boundaries such as state/society, nature/culture, and social/scientific, ANT proposes an ontology in which nonhumans and humans are analytically symmetrical in collective assemblages. It is these collective assemblages, in their complex and material relationality, that ought to be explored. ANT's concepts of heterogeneity, material relationality, and punctualization, lead to a *distributed* conception of social action concerned with a multiplicity of human and nonhuman actants. Such a thesis not only moves ANT closer to what feminist STS has considered *post-humanism*, it also orients its analytic tools to the ontological complexity of action in the world.

Sam Haraway and Patrick Carroll

See also Pragmatism; Social Construction of Scientific Knowledge

Further Readings

- Callon, M. (1986). Some elements of a sociology of translation: Domestication of the scallops and fishermen of St. Brieuc Bay. In J. Law (Ed.), *Power, action and belief: A new sociology of knowledge?* (pp. 196–223). Routledge.
- Carroll, P. (2012). Water and technoscientific state formation in California. *Social Studies of Science*, 42(4), 489–516.
- Latour, B. (1988). *The pasteurization of France*. Harvard University Press.
- Law, J. (1992). Notes on the theory of the actor network: Ordering, strategy and heterogeneity. *Systems Practice*, 5(4), 379–393.
- Pickering, A. (2008). New ontologies. In A. Pickering & K. Guzik (Eds.), *The mangle in practice: Science, society, and becoming* (pp. 1–15). Duke University Press.

AD HOC HYPOTHESIS

The topic of this entry, the ad hoc hypothesis, is a small and recent footnote in a long tradition of debating how best to determine scientific truth.

On the surface, the concept is straightforward: a hypothesis is formed *ad hoc*, that is, in response to a particular situation instead of in anticipation of a general application. A host of validity issues are prompted by the introduction of an explanation ad hoc: validity of the theory, of the investigation, of the findings, and of the generalizations that may follow. These issues are the context in which the term must be understood, beginning with the fact that it is usually associated with pseudoscience, that is, conclusions which cannot be verified independently. That said, it is important to recognize a new context in which ad hoc hypotheses figure: belief revision within computer programming for artificial intelligence.

The term itself has prompted great debate concerning its definition, its purpose, and its value. Let us first consider the term *ad hoc*, a Latin phrase meaning *for this*, which is used in reference to something that is *one-off*, that is, not an established and predictable routine, but formed for a specific and short-term purpose. An ad hoc committee functions as a task force with a brief to address one issue, and then it is disbanded once the issue is addressed. The colonial militias were ad hoc because they formed in response to an immediate threat after which the soldiers reverted to their chosen occupations. In contrast, a standing committee or a standing army will have well-defined continuous roles in relation to the agencies they serve, and their members develop more professional expertise, whereas ad hoc groups operate on the common sense and skills of the average agency member. This adds to the more casual association attached to *ad hoc hypotheses*, which are also considered less rigorously developed and less valuable. An ad hoc hypothesis is produced to account for anomalies that are inconsistent with the original hypothesis.

Hypotheses are propositions of general truth, that is, a prediction of an outcome that will naturally occur based on the logic of a theory or belief. These statements are formal interpretations of the hunches and guesses that prompt researchers to investigate phenomena. Hypotheses are used to limit the scope of the investigation that is intended to test the validity of the theory prompting the research question. If the findings are inconsistent with the predictions of the theory, the theory itself is then in doubt. The value of a theory lies in its

capacity to explain naturally occurring phenomena and to help people predict and perhaps control future events. This reasoning is based on a positivist orientation to reality, meaning that the world is understood to be fairly stable and, given enough rigorous observation, its patterns can be recognized with some confidence and in turn used with confidence in similar situations. At issue is the presence of anomalies that are inconsistent with the theorized expectations, and whether they should be used to discard the theory or inform the theory. Empirical investigations are designed to test hypotheses in a systematic way so that others can replicate the procedure and thereby test the validity of the findings. Adding weight to this debate is the pride of integrity that prompts many researchers to distance themselves from practices inconsistent with their epistemological orientation and even to disparage those who appear to compromise the quality of rigor they seek.

It is safe to say that there is a keen competition among academics to claim authority. A theory is considered less robust if it cannot withstand general application without special adjustments to different circumstances. Whether termed ad hoc hypotheses or rival hypotheses or counterarguments, statements are considered fallacies if they are based on factors that cannot be tested, such as paranormal activity, or *fate*, or superstition. They may be dismissed as simply unhelpful because they distract from the task at hand or may be found unattractive because they disregard the elegance of a simple equation. They rarely add persuasive plausibility. As found in the political arena, ad hoc explanations are called *spin control* and are associated with a cynical manipulation of language to camouflage the failure of the findings to be self-evident. Within a peer review of an investigator's report, there will be keen criticism if the discussion introduces theory that was not established in the literature review justifying the research hypothesis. There is instead a convention of intellectual humility: to identify future research trajectories given the limitations of the study.

In the larger context of advancing scientific knowledge, however, the practice can be viewed as developmental in the professional researcher. The rigidity of a novice's insistence on the absolutes of binary opposites gives way to the flexibility of an expert who can tolerate ambiguity, paradox, and

multiple explanations in the process of evaluating the merits of a theory. The naïve and inexperienced are motivated to defend a preconceived notion, while the more experienced will consider the pattern of empirical evidence. These are in fact the argumentational and rhetorical skills necessary for scientific discourse: generating evidence and counterarguments to establish proof and certainty. Ad hoc hypotheses are a function not of the original experimental design but of its interpretation by the investigator. The investigator may believe that this modification improves the explanation of the phenomenon, but the more traditional scientific community tends to regard this as an evasion, motivated by a desire to protect the theory in question from disputation.

The role of ad hoc hypotheses could be absorbed into the fabric of basic scientific discourse were it not for a current interest in mechanizing the executive function of decision. *Artificial intelligence* is a contemporary field attempting to program computers with all possible contingencies so that the program can actually *learn*; that is, revise its own procedures based on responses. A computer program is a series of propositions requiring the programmer to predict the nature of decisions to be made, contingent on other actions. The algorithm by which the program proceeds through these decisions is of course crucial to the outcome, for a simple reversal of steps can result in a significant change. *Belief revision* in this context does not mean a change in philosophical conviction about reality or morality, but a change in technological procedures that are, from the standpoint of the machine, the way things are. These are fundamental issues for the organization of information within all fields, so a belief revision theory launched a wide-ranging field of research. This theory is found in the Alchourrón, Gärdenfors and Makinson (AGM) model, a formal framework for analyzing ways to change beliefs in the context of computer programming. The three researchers collaborated from different fields to propose it, illustrating the multidisciplinary interest in this logic. Complex mathematical formulae are the currency of this discourse, akin to setting up an intricate series of dominoes in readiness for a simple action to trigger a long sequence of response. This enterprise assumes a closed system of variables; that is, it operates within a set of propositions, but as

databases accumulate, they must be updated, leading to an important concept of database priorities within truth maintenance systems. They are also deterministic in that given a set of beliefs responding to some input, the resulting belief set is well anticipated. A combination of operators may be used as an indeterministic operator that allow more than one admissible outcome. Ad hoc hypotheses are thus a mainstay of the dynamic process.

Ad hoc hypotheses highlight the epistemological distinctions between belief and knowledge, between intuition and theory, and between deductive and inductive approaches to problem-solving. Opinions about what ad hoc hypotheses are, and their value, reflect one's orientation to identifying truth based on observed phenomena, on principles derived from an accepted theory; for example, syllogisms associated with Aristotle, or on authoritative conclusions such as those demonstrated in Socratic dialogues, which used hypothetical situations for support of interpreting ideals. Modern rationalism follows René Descartes' four steps of scientific thinking: skepticism (that nothing is true until proven so), reductionism (divide concepts into much smaller entities that are more easily observed), organization (ordering knowledge from simple to complex), and generalization (refining a theory to eliminate omissions and contradictions). Francis Bacon argued that theory must be built from the ground up by first accumulating a large sample of data, then proposing some preliminary axioms or predictable patterns, and finally generating a theory that would explain all the phenomena.

Naomi Jeffery Petersen

See also Falsifiability; Justification; Refutation; Scientific Realism

Further Readings

- Bamford, G. (1999). What is the problem of ad hoc hypotheses? *Science and Education*, 8(4), 375–386.
- Bamford, G. (1993). Popper's explications of 'Ad Hoc'ness: Circularity, empirical content, and scientific practice. *British Journal for the Philosophy of Science*, 44(2), 335–355.
- Hunt, J. C. (2012). On ad hoc hypotheses. *Philosophy of Science*, 79(1), 1–14.

Klein, E. E., & Herskovitz, P. J. (2007). Philosophy of science underpinnings of prototype validation: Popper vs. Quine. *Information Systems Journal*, 17(1), 111–132.

Patton, L. (2012). Experiment and theory building. *Synthese*, 184(3), 235–246.

ANALYSIS

According to the prevailing definition, analysis is a process that involves breaking something up into simpler or component parts. Indeed, the word *analysis* derives from the Greek *ana* for up and *lysis* for loosening or separation. But analysis can also be conceived as the search to determine prior principles, beginning from what is known—a conception with roots in ancient geometry and with purchase in academic philosophy. Analysis can be contrasted with its converse, synthesis, which is the process of joining parts or elements together to create a more unified result. Both synthesis and analysis are directed toward the ends of elucidation; through analysis, that which is being analyzed becomes better or more fully understood.

Analysis is a fundamental mode of inquiry and is applicable without restriction; one can conceivably analyze anything. As a result, it occurs throughout lay and specialized academic contexts. In their daily lives, people regularly engage in analysis, analyzing the texts they read or watch, the events they attend, the words, deeds, and behaviors of other people, and so on, discerning component parts of larger wholes. Similarly, scholars employ analysis across a wide range of academic contexts, including the sciences, social sciences, arts, business fields, and in math and the humanities, where it is a basic building block of intellectual pursuit. Scholarly analysis is generally distinguished from lay analysis by a greater degree of rigor and formalization, though the extent to which it is either depends on the particular discipline and method.

As a decompositional process, analysis includes both descriptive and pattern-seeking activities. As descriptive, analysis involves identifying the parts of a whole and providing an account of the relevant features of those parts, in some greater or lesser degree of detail. As a result, it is closely tied

to observation; and the observation that drives analysis can involve any of the senses: visual, auditory, gustatory, tactile, or olfactory. Moreover, in order to facilitate analysis or, in some cases, even make it at all possible, one might utilize instruments to augment one or more of the senses. The analysis of physical entities or qualities, for examples, often involves the use of instruments or tools. But perhaps more often than being purely descriptive, analysis also entails the search for patterns. Breaking a larger and more complex whole into smaller and more elementary parts brings the interactions between those parts into focus. And in analyzing, one classifies and groups, identifying relationships and drawing conclusions about them.

Analysis is a comprehensive term that refers to a highly generalized activity, which can be actualized in innumerable different ways. For example, analysis can proceed at any level of abstraction. While a biologist might probe a tangible organism in order to analyze its concrete parts, philosophers use conceptual analysis to examine thoroughly abstract concepts like *truth* and *knowledge*. Levels of granularity also vary widely; coarse-grained macro-analyses assess relatively large-scale components, while fine-grained micro-analyses consider relatively small ones. One can even analyze analyses, as is the case in a meta-analysis. Indeed, there are innumerable different approaches and/or specific formalized methods one can employ in order to analyze a target of interest, some of which are outlined in the subsequent sections of this entry.

Analysis is an inherently interpretive activity. In the processes of analyzing, one makes decisions about where to direct attention, about what constitute the meaningful parts of a whole, and construes the existence of relationships and patterns. And while such judgments are sometimes relatively clear-cut, they can also be uncertain and arguable. In the act of breaking something down into its component parts, therefore, we explain how to understand it and make a case for what it means. We render it otherwise. As a result, the terms *analysis* and *interpretation* are sometimes used synonymously. For example, an analysis of a poem might also be described as an interpretation of it. Interpretation, however, does not have the decompositional emphasis that analysis does.

Moreover, interpretation often tackles questions of meaning that fall outside of the purview of analysis. Similarly, analysis differs from criticism, another related intellectual pursuit, in that it is not in essence evaluative.

An analysis is not only the act of breaking something down into component parts but also the communicated record or result of so doing. The form of that resultant analysis depends on the type of analysis one has done and the type of context in which one has performed it.

Deductive Approaches

In deductive approaches to analysis, an existing hypothesis, theoretical model, or conceptual framework guides the descriptive and pattern-seeking activities of analysis. Therefore, the hypothesis, theory, or framework directs attention, determining to some extent the parts, patterns, and relationships delineated. As such, they can be characterized as offering a perspectival lens. For example, in a hypothetico-deductive experimental paradigm, one analyzes data in relation to a preidentified hypothesis to determine whether or not it supports that hypothesis. Similarly, a deductive analysis of a dream guided by Freudianism will focus on features specified by or emphasized in Sigmund Freud's psychoanalytic theory, such as repressed memories and unconscious desires.

Within disciplines, certain conceptual frameworks tend to predominate, becoming frequently engaged as analytic tools. In the humanities and social sciences, for example, psychoanalysis, Marxism, feminist theory, structuralism, critical theory, phenomenology, and deconstruction have all been commonly employed toward the ends of deductive analysis. Those employing these and other models in a deductive fashion must guard against confirmation bias, or the tendency to only look for/see that which confirms an existing hypothesis, overlooking those features that do not fit within its parameters. Similarly, deductive approaches are vulnerable to what is sometimes called *cooker-cutter criticism*, whereby a theoretical framework so strongly and narrowly governs analysis that it becomes a rote and mechanical application of theory with predetermined and not particularly insightful results.

Although an explicitly acknowledged and articulated theoretical framework drives some analyses, the principles that guide analysis in a deductive fashion are sometimes implicit. This is particularly true in lay contexts, where neither theoretical frameworks nor the process of analysis are typically formalized, but also occurs in scholarly contexts as well. Moreover, folk theories sometimes deductively underlie scholarly analysis in a covert fashion. The mobilization of implicit theoretical frameworks can be problematic given that the premises by which the analysis is guided are not articulated and may not be apparent to the person doing the analysis and/or the person encountering the results of the analysis, making them unavailable for discussion or critique.

Inductive Approaches

In inductive analysis, inquiry begins with the analysandum itself; one works with and from the object of interest in order to discern its meaningful parts and draw conclusions about the relationships between those parts. That is, the *breaking down* of this form of analysis involves a *building up* from particulars as one moves from specific features to more general conclusions about parts and patterns. An example of a commonly employed inductive method of analysis in the humanities is close textual analysis, which involves paying scrupulous attention to the details of a selected work or works, from the use of commas to the overarching organization, to identify patterns within that work and/or among works.

However, though theory does not drive an inductive analysis, it must inevitably be guided by something, which often includes intuition or past experience with other texts. This raises the question of the extent to which any analysis can in fact shelve theory, and whether inductive approaches are in fact inevitably theory-laden, even if that theory is of the informal and implicit kind that comprises intuition and develops out of experience. Moreover, philosophical debates about and psychological investigations into the theory-ladenness of observation are clearly applicable to analysis given its reliance on observation.

The human impulse to see and seek patterns and connections not only results in a general affinity for analysis, but also a propensity to see

patterns and perceive connections where there are none. This tendency underlies many superstitious and supernatural beliefs and is responsible for much of what is termed *magical thinking*, but it is also a tendency with implications for scholarly analysis. *Overreading* or *overinterpretation* are terms sometimes used to describe overly aggressive acts of analysis in which patterns are read into a text rather than being found there; and this problematic analytic misstep can be found in all sorts of analyses, from the poetic to the radiographic.

While deduction and induction are distinguishable orientations toward analysis, in practice the two often co-occur. They do so in a compelling way in the *grounded theory method*, a formalized methodology developed in the social sciences by which one self-consciously and systematically puts inductive analysis toward the ends of theory development. In using analysis to arrive at theory, one reverses the typical or traditional pattern by which theory precedes analysis. It is an iterative process in that one repeatedly moves back and forth between data and analysis, and between inductive and deductive modes of thought, becoming increasingly abstract and theoretical as the process unfolds.

Quantitative Analysis

Quantitative analysis is the analysis of that which has been quantified; that is, calculated, counted, or measured. And in practice, it can take many different forms. Perhaps the simplest type of quantitative analysis is the discernment of relative quantities. A quantitative chemical analysis of this type, for instance, might determine the amounts of the various component parts of a more complex chemical. But quantitative analysis often moves beyond the relatively straightforward assessment of amounts, to discern patterns and relationships, which is a way of making numerical data meaningful and/or useful. Quantitative analysis can also be put toward the ends of prediction. In predictive analysis, one looks at what has happened in the past and identifies patterns in order to predict future events, as in analyzing voting behavior in previous election cycles to predict outcomes in upcoming ones.

Statistical methods are mathematical tools for seeking out and delineating the patterns in a

quantified subject of analysis. Indeed, statistics, a branch of mathematics, is the science of quantitative analysis. Statistics can be applied in a deductive fashion, as in confirmatory data analysis, whereby one identifies a hypothesis and employs statistical analysis in order to test it. Or one can use statistics in a more inductive fashion, applying them to a set of data in order to see what patterns emerge, a type of exploratory data analysis. There are two main branches of statistics, descriptive and inferential. *Descriptive statistics* are methods for characterizing the features of a given set of data; these include measures of central tendency (i.e., mean, median, and mode) and spread (i.e., range, variance, and deviation). *Inferential statistics* are methods for using a sample in order to identify the parts, patterns, and relationships of a larger population; in doing so, the quality of the sampling process is key. *Regression analysis* also deserves mention as a set of powerful statistical processes for examining the relationships between variables, often independent and dependent ones.

In some quantitative studies, particularly those in the social sciences, coding is an important initial step in the analysis process. It is the procedure by which one applies shorthand labels to data items, such as parts of interview transcripts or field notes; in doing so, one not only categorizes them and enables them to be more readily sorted, but also makes them countable and thereby amenable to analysis via quantitative techniques. For example, content analysis, a technique for analyzing communicative acts, often employs coding toward the ends of quantification. When coding, one must decide on the unit of analysis (e.g., words, paragraphs, parts of images, entire photographs) and whether to use predetermined codes or to derive codes during the course of coding. The process of coding data does some analytical work in that it discerns recurring features. It also facilitates subsequent *higher-level* analysis. The reliability of the coding process is an important consideration, and studies often consider *intra-coder reliability*, or the consistency with which a given person codes the data, and to enhance reliability, some studies employ multiple coders, and consider *inter-coder reliability*, or the extent to which they agree with each other.

Visualization is also an integral part of quantitative analysis. The results of quantitative analysis

are often effectively presented in visual form, which can incisively convey patterns and relationships between component parts. Tables and graphs, frequency distributions, bar charts, scatterplots, and histograms are all commonly used to present the results of a quantitative analysis. Moreover, data visualization is itself considered a method of quantitative analysis and a branch of descriptive statistics. As a method of analysis, we can consider the process of visualization itself as that which discerns meaningful components and their relationships.

As a result of advances in computing, quantitative analysis has become not only more large-scale and sophisticated, but also increasingly democratized. While the computing power necessary to do complex statistical analysis used to be in the hands of only a few, it is now much more widely available. Moreover, there is now unprecedented access to vast data sets, such as those generated via new media, offering opportunities large-scale quantitative analysis with considerable potential across academic disciplines and for commercial purposes. In professional fields, the word *analytics* is often used to refer to the science of data analysis.

Qualitative Analysis

Qualitative analysis is the analysis of that which has not been (and perhaps cannot be) measured or quantified. It focuses on qualities and features rather than amounts. For instance, a qualitative chemical analysis might detail the characteristics of the component parts of a more complex compound. As with quantitative analysis, qualitative analysis can proceed inductively or deductively; that is, one can take the detailed particulars of the analysandum as the analytic starting point or can proceed on the basis of a theoretical framework.

While the phrase *qualitative analysis* might well be used to identify a wide range of analytic inquiry, it is generally associated with a research paradigm in social science fields that emerged during the 20th century to develop and systematize qualitative methods as legitimate alternatives to positivistic and quantitative ones. As a result, not all of those who engage in analysis that could certainly be described as qualitative would necessarily use that phrase to describe

their work, particularly those who work in the humanities. For example, philosophy's tradition of conceptual analysis, which is solidified in the branch of philosophy known as *analytic philosophy*, is certainly not quantitative, but philosophers would not tend to identify it under the banners of *qualitative analysis* and *qualitative research*.

In the social sciences, qualitative analysis tends to be rather systematized and is often part of a research process with parallels to scientific research. For example, the qualitative data are often systematically generated or collected for the purposes of analysis; this is in contrast to the analysis of something preexisting and simply selected for study, as in a rhetorical analysis of an influential speech or the art historian's analysis of a famous painting. Qualitative researchers in the social sciences might conduct interviews, run focus groups, and engage in participant observation, analyzing the transcripts or notes that result from these activities. These textual data are also often coded, a process that is not only an instrument of quantification but is can be a means of sorting data and thereby facilitating qualitative analysis. Some of the specific approaches to qualitative analysis that are widely used in the social sciences include narrative analysis, discourse analysis, phenomenological analysis, and grounded theory.

Qualitative and quantitative analysis can be complementary ways of discerning the parts and patterns of a subject of interest, and many research projects, such as mixed-methods studies, use both. Moreover, social scientists sometimes use qualitative analysis in order to engage in initial exploration of a subject, followed by quantitative analysis to test hypotheses derived from those exploratory qualitative endeavors.

Critiques of Analysis

Alongside the widespread embrace of analysis as fundamental mode of intellectual engagement exists a strain of skepticism regarding its power to elucidate. The so-called *paradox of analysis*, associated with C. H. Langford and G. E. Moore, is, essentially, the idea that to give a correct analysis of a concept, one already has to know and be able to identify its component parts. For instance,

Moore provides the example of analyzing the concept of brother into male and sibling, something that requires already knowing what a brother is. The analysis is trivial and does not provide any new insights. As a result, it follows that it is impossible for an analysis both to be correct and to inform us of something we do not already know. And although debates regarding paradox of analysis (which include attempts to resolve it) occur specifically in relation to philosophical conceptual analysis, the paradox has relevance to analysis more broadly considered.

Others express concern regarding analysis's essential reductiveness; by definition, analysis involves breaking things down, distilling the elementary parts from the more complex whole. Philosophers such as Georg Wilhelm Friedrich Hegel and Friedrich Schiller have pointed toward the reductive dimension of analysis. Pierre Teilhard de Chardin once described analysis as "that marvelous instrument of scientific research to which we owe all our advances but which, breaking down synthesis after synthesis, allows one soul after another to escape, leaving us confronted with a pile of dismantled machinery, and evanescent particles" (1959, p. 283). That is, the elucidation of analytic dismantlement is not without its losses.

Michelle G. Gibbons

See also Abstraction; Conceptual Analysis; Generalization; Interpretation

Further Readings

- Beaney, M. (2014). Analysis. In E. N. Zalta (Ed.), *The Stanford encyclopedia of philosophy* (Summer 2014 ed.). Retrieved July 2014, from <http://plato.stanford.edu/archives/sum2014/entries/analysis>
- Berelson, B. (1952). *Content analysis in communication research*. Free Press.
- Glaser, B., & Strauss, A. (1967). *The discovery of grounded theory: Strategies for qualitative research*. Aldine.
- King, J. C. (1998). What is a philosophical analysis. *Philosophical Studies*, 90, 155–179.
- Miles, M., Huberman, B., Michael, A., & Saldaña, J. (2014). *Qualitative data analysis: A methods sourcebook* (3rd ed.). SAGE.
- Selzer, J. (2004). Rhetorical analysis: Understanding how texts persuade readers. In C. Bazerman & P. Prior

(Ed.), *What writing does and how it does it: An introduction to analyzing texts and textual practices* (pp. 279–308). Erlbaum.

Teilhard de Chardin, P. (1959). *The phenomenon of man* (Trans. B. Wall). Harper.

Tukey, J. W. (1977). *Exploratory data analysis*. Addison-Wesley.

APPLICABILITY OF MATHEMATICS IN SCIENCE

Our understanding of nature, our explanations of natural phenomena, crucially depend on the use and interpretation of mathematics in scientific practice. Mathematics provides a common language for the empirical (i.e., the natural and social) sciences, and through its tools, we get successful predictions and empirical confirmations that are about the entities and the phenomena of these sciences. When this happens, we say that mathematics successfully applies in science. But the success of mathematics in science prompts the following question: How is it possible for the abstract entities of mathematics (numbers, equations, integrals, etc.) to say something interesting about phenomena and entities, as for instance nuclear reactions and proteins, which are studied by the empirical sciences and have a completely different nature? The issue is philosophical, and the research topic associated to it is known as the *Applicability of Mathematics in Science*. Following a short introduction to the topic, this entry offers an overview of some accounts that have been proposed to tackle the philosophical issue. It also identifies some connections that have been established between the study of the applicability of mathematics in science and other debates in philosophy of science and philosophy of mathematics.

The Philosophical Challenge

When will Mars be closest to Earth? What is the distance covered by a falling body in a given time under the force of gravity? Given a protein with a specific amino acid sequence, what will its three-dimensional structure be like? What is the average

number of predators in a particular ecosystem? How can we maximize profit in a market economy with production constraints? Although these questions involve entities (planets, falling bodies, proteins, predators, production constraints) and empirical sciences (physics, computational biology, ecology, economics) that are quite different from each other, there is a common path to answer them: Do the math! Surprisingly, mathematics works. This, of course, does not mean that mathematics always works when it is used, or *applied*, in the empirical sciences. Nevertheless, in many cases, mathematics provides a correct, empirically accurate, answer to our original question about planets, proteins, predators, and so on. Why should we be surprised about such effectiveness, or successful applicability, of mathematics? The issue is philosophical: Mathematical entities are generally regarded as abstract because they do not causally interact with anything (e.g., a number cannot break a glass) and have no spatiotemporal location (e.g., a number cannot be found in a certain place at a certain time); therefore, it is quite surprising that they can properly inform us on the behavior and properties of entities like planets or proteins, which are causally active and are located in space and time. Even more surprisingly, sometimes the mathematics that is successfully used in science is not developed from scratch to study an empirical problem, but it has been developed within a purely mathematical framework before being used with success in a specific context of application (the mathematics of group theory, for instance, was developed long before its introduction into quantum physics in the late 1920s).

It is exactly this particular feature of mathematics, namely that of being applicable with success outside its boundaries, that allows scientists to make extraordinary progress in the understanding of nature. And the main goal of philosophers of science and mathematics working on the applicability of mathematics in science is to investigate such features and account for the (successful) application of mathematics in science.

Facing the Challenge

The applicability of mathematics has a long-standing philosophical pedigree. Indeed, the first reflections on the successful use of mathematics to

understand the world can be traced back at least to early Pythagoreanism (sixth through fourth centuries B.C.E.) and to the investigation of harmonic intervals in terms of arithmetic. In classical Greece, harmonics was the science that investigated the arrangements of pitched sounds that form the basis of musical melody and the principles that govern them. What the early Pythagoreans discovered is that some differences in pitch between two sounds, or intervals, were expressible by numerical ratios (e.g., they found that the three basic intervals of octave, fifth, and fourth were expressible as the numerical ratios 2:1, 3:2, and 4:3). Thus, an arithmetical ratio could be successfully used to represent a perceptible phenomenon. To account for this fact, they proposed a metaphysical system wherein everything is a number understood via discrete representations. In this “all is number” doctrine, which is illustrated in Aristotle’s *Metaphysics*, the world consists of corporeal numbers, or units, and mathematics applies to the world because the world has such mathematical structure.

Although the Pythagoreans pioneered the study of our topic, later attempts to explain the applicability of mathematics in science departed from their view. The Greek philosophers Plato (ca. 429–347 B.C.E.) and Aristotle (384–322 B.C.E.), for instance, developed different approaches to the applicability issue.

For Plato, the physical world and all of its constituents are imperfect copies, approximations, of special abstract entities called *Forms*. The world and its constituents derive from Forms, which include—but are not limited to—mathematical entities and properties. For instance, a drawn square belongs to the material world, but it is an imperfect, approximate copy of the square as such, whose angles are exactly 90° and whose sides are perfectly straight. Because all copies are dependent on ideal Forms, the physical world is dependent on Forms. Therefore, for Plato, mathematics is applicable to the physical world because it is a science of the ideal Forms from which the physical world is derived.

Aristotle disagreed with Plato about the nature of mathematical objects. While for Plato material objects are imperfect copies of ideal forms, and the former are derived from the latter, Aristotle maintains that mathematical objects are obtained

from physical objects through a process of *abstraction*: We consider a physical object with respect to some of its attributes, next we take them out, and finally we create a new (abstract) object consisting in that object in only those respects. The result of this process is an object in its own right. The *snub-nose* is Aristotle’s favorite example: If we consider a snub nose, the quality of snubness comes from sense perceptions and is concavity in the physical object *nose*; the mathematical quality of concavity, on the other hand, is obtained by removing all the attributes of the nose with the exception of its shape. Thus, according to Aristotle, mathematics studies the objects formed by abstraction but not the material objects on which the abstraction process operates. And this view provides an account of how mathematics applies to the world.

The philosophical standpoints proposed by Plato and Aristotle were very influential until the early modern period. The Italian astronomer, engineer, and physicist Galileo Galilei (1564–1642), for instance, echoes both views in *The Assayer* (1623) and in the *Dialogue Concerning the Two Chief World Systems* (1632). But the mathematization of physics that took place during the 17th century, largely influenced by the publication of Isaac Newton’s *Principia* (1687), marked a radical departure from the dominant tradition of a natural philosophy that aimed to give qualitative explanations of natural phenomena. In the new conceptual framework that arose, where the objects of physics also included objects that were themselves mathematical (e.g., Newton’s concept of force), philosophical analysis such as those offered by the Pythagoreans, Plato, and Aristotle were insufficient to account for the applicability of mathematics. Therefore, philosophers had to elaborate new strategies to explain the applicability of mathematics. Immanuel Kant (1724–1804), for instance, proposed—within his transcendental idealism—the idea that it is the constitution of the subject that makes it necessary to describe the world mathematically. In his *System of Logic* (1843), John Stuart Mill argued that mathematics is itself empirical since mathematical truths are generalizations derived from experience. Other philosophical standpoints were offered. Nevertheless, as happened with the accounts proposed by the ancient Greeks, these conceptions had to be revised or even abandoned in light of new, unexpected interactions between

mathematics and science (e.g., the application of non-Euclidean geometry in Albert Einstein's theory of general relativity undermined Kant's view, and particularly his conception of the synthetic *a priori* character of geometry).

Successful interactions between mathematics and the empirical sciences have been growing massively since the first half of the 20th century. This phenomenon captured the attention of many scientists, as for instance the Nobel laureates in physics Eugene Wigner and Richard Feynman, and gave a new impetus to the philosophical study of the applicability of mathematics. Several analyses were developed but, in general, we can identify three main strategies that have been adopted to address the applicability issue: (1) a logicist approach, (2) an attempt to explain the applicability of mathematics by focusing on the way in which mathematical concepts are introduced in scientific theory-building, and (3) a strategy that elaborates on the idea that the applicability of mathematics can be explained in terms of some correspondences that can be established between a mathematical domain and an empirical domain. Let us consider the three strategies in turn:

1. In philosophy of mathematics, *logicism* is the view that mathematics, or part of it, can be reduced to logic. This view was developed by the philosopher Gottlob Frege at the end of the 19th century. According to the logicist standpoint, mathematics can be applied to the description of the world because it contains only logical concepts, which can be applied to any possible domain of objects. Although Frege's attempt to demonstrate logicism was undermined (because of an inconsistency that was found in his formal logical system), by the mid-1960s, some philosophers of mathematics tried to revive some of the core ideas of logicism and set in motion a philosophical movement that has been called *neologicism*. Neologicists defend *Frege's constraint* (also known as *application constraint*), a principle that was one of the main tenets of Frege's logicism and which requires that the empirical applications of a mathematical theory (or class of numbers) must be built directly into the formal characterization of the mathematical theory (or class of numbers). Thus, although in a renewed form, neologicist projects that meet Frege's constraint continue to offer an account of the applicability of mathematics by demanding that an explanation of applicability be provided inside mathematics, by mathematical definitions.
2. A second strategy, which was largely inspired by the emergence in the 1960s of a tradition in philosophy of mathematics that was concerned with the dynamics of mathematical discovery and the historical development of mathematics, focuses on the way in which the introduction of mathematical concepts is influenced by theory-building in science. Philosophers who follow this route draw on analyses of the historical development of mathematics and present case studies in which mathematics and science simultaneously develop in a theory-building context. By emphasizing the interplay between mathematics and other kinds of practices that we find in theory-building (e.g., the practice of identifying analogies between different scientific theories), they argue that the effectiveness of mathematics in science is no longer surprising when we see that (and how) much mathematics is brought into being by the need to model one or more aspects of the world (an example of such an analysis can be found in the paper "Solving Wigner's mystery," written by the historian of mathematics and logic Ivor Grattan-Guinness and published in 2008).
3. A third strategy, which has been the most influential since the first decade of the 21st century, gives an explanation of the applicability of mathematics in terms of mappings that are established between mathematics and the empirical systems we want to study. These mappings are mathematical and include not only isomorphisms but also other kinds of mathematical mappings like homomorphisms, epimorphisms, and monomorphisms. Mappings ensure that some crucial features of the empirical system are mirrored in the mathematical model used to study that system. Such an approach has been called *the mapping account of applied mathematics*. According to this view, the applicability of mathematics is fully accounted for by appreciating the relevant structural similarities between the empirical system under study and the mathematics used in the investigation of that system.

There have been many attempts to implement the mapping account, but the most discussed has been that proposed by Otávio Bueno and Mark Colyvan in their 2011 paper “An inferential conception of the application of mathematics.” Bueno and Colyvan retain the basic idea behind the mapping account (i.e., the idea that the successful application of mathematics amounts to establishing a suitable mapping between mathematics and an empirical setup) and propose a view of applicability, called *the inferential conception*, in terms of a three-step process: (a) immersion (a correspondence between mathematics and the empirical setup is established via a suitable mapping), (b) derivation (some consequences are generated from the mathematical formalism), and (c) interpretation (the mathematical consequences obtained in the derivation step are interpreted in terms of the initial empirical setup, via a mapping that does not necessarily coincide with the mapping used in the immersion step).

Connections With Other Debates

The growth in successful interactions between mathematics and science played a motivating role in advancing analysis of the applicability issue. Nevertheless, philosophical interest in the applicability of mathematics was also prompted by the investigation of three topics that have become increasingly central to the agenda of philosophers of science and mathematics since the 1970s: mathematical modeling in science, mathematical explanation in science, and indispensability argument(s) for mathematical Platonism.

Mathematical Modeling in Science

Mathematical models are used in science to represent and study a selected part or aspect of the world, which is named the (model’s) *target system*. For instance, the so-called Lotka–Volterra model, developed by the mathematicians Alfred Lotka and Vito Volterra in the 1920s and consisting in a pair of differential equations, is used in population ecology to represent a system of two interacting species (a predator and a prey) and study its dynamics. Sometimes, a mathematical model can also be used to represent and study two different target systems (e.g., the same equation can be used

to represent and study a swinging pendulum and an oscillating electric circuit). Furthermore, the target system represented in the model is usually an *idealized* version of the real system, namely a simplified or distorted version of it. For instance, when we study a swinging pendulum, we usually assume that the pendulum bob is a point mass and that the string is massless. With these idealizations in place, the target system becomes more tractable from a mathematical point of view.

It is therefore easy to see how the investigation of the applicability of mathematics is closely connected to discussions of mathematical modeling in science: An account of how a mathematical model can be used to represent and study a target system prompts an explanation of how mathematics is applied with success to the physical world (or, if we consider the case of idealizations, an explanation of how mathematics is applied with success to those idealizations that are about the physical world).

Mathematical Explanation in Science

To explain something is to give the reason why that something happens (rather than a mere description of it). For instance, to explain a solar eclipse is to give an answer to the question “Why did the solar eclipse occur?” rather than a simple description of the eclipse phenomenon. Explanations are therefore particularly important in science, which investigates the reasons why empirical phenomena (such as the solar eclipse) occur.

Many explanations in science appeal to causes. These *causal explanations* provide information about what caused the thing that we want to explain (e.g., a solar eclipse is caused by the interposition of the moon between the sun and the earth). Moreover, many scientists and philosophers have argued that there also exist mathematical (and therefore noncausal) explanations in science. In this kind of explanation, it is mathematics that unveils the reason why an empirical phenomenon occurs. Nevertheless, there could be no mathematical explanation of empirical phenomena if mathematics were not successfully applied in the sciences dealing with such phenomena (note that the converse of this conditional statement is not true since there may be successful applications of mathematics that are not

explanations). Therefore, articulating a plausible account of the applicability of mathematics in the empirical sciences is particularly important for those philosophers who maintain that mathematics can play an explanatory role in science.

Indispensability Argument(s) for Mathematical Platonism

Platonism in philosophy of mathematics is the view that mathematical objects exist (independently of us and our language, thought, and practices). The indispensability argument is a philosophical argument used by Platonists to defend their view and argue that there are mathematical objects. There exist several versions of the argument but, in its original formulation (which is credited to the philosophers Willard Van Orman Quine and Hilary Putnam), the argument goes as follows: Since we ought to believe in the existence of those entities that are indispensable to our best scientific theories (like electrons or black holes), and since mathematical entities are indispensable to our best scientific theories (i.e., our best scientific theories like quantum mechanics and general relativity cannot be formulated without making reference to mathematical entities), then we ought to believe in the existence of mathematical entities. In this line of argument, mathematical entities are seen to be on an epistemic par with other theoretical entities of science like electrons or black holes (and so anyone who believes in the existence of the former should also believe in the existence of the latter). Furthermore, the argument crucially presupposes that mathematics successfully applies in science. Hence, again, we see how investigations into the applicability of mathematics are linked to, and triggered by, the study of a separate topic.

Although discussing other formulations of the indispensability argument is beyond the scope of this entry, it is worth noting that there is a version of the argument that relies on the notion of mathematical explanation in science. Early presentations of such an *enhanced* (or *explanatory*) indispensability argument can be found in philosophy of mathematics before the end of the 20th century, but an explicit formulation appears in 2005 in Alan Baker's paper "Are there genuine mathematical explanations of physical phenomena?" Since, as

discussed earlier in this entry, successful applicability of mathematics is a necessary—although not sufficient—condition for mathematical explanation in science, the issue of explaining why mathematics applies with success in science also arises in the context of the enhanced indispensability argument.

Daniele Molinini

See also Explanation; Inference to the Best Explanation; Laws (Scientific); Mathematics, 20th Century; Modeling; Philosophy of Science; Physics, 20th Century; Realism in Mathematics

Further Readings

- Baker, A. (2005). Are there genuine mathematical explanations of physical phenomena? *Mind*, 114(454), 223–238. <https://doi.org/10.1093/mind/fzi223>
- Bangu, S. (2012). *The applicability of mathematics in science: Indispensability and ontology*. Palgrave Macmillan.
- Bueno, O., & Colyvan, M. (2011). An inferential conception of the application of mathematics. *Noûs*, 45(2), 345–374. <https://doi.org/10.1111/j.1468-0068.2010.00772.x>
- Bueno, O., & French, S. (2018). *Applying mathematics: Immersion, inference, interpretation*. Oxford University Press.
- Gingras, Y. (2001). What did mathematics do to physics? *History of Science*, 39(4), 383–416. <https://doi.org/10.1177%2F007327530103900401>
- Grattan-Guinness, I. (2008). Solving Wigner's mystery: The reasonable (though perhaps limited) effectiveness of mathematics in the natural sciences. *The Mathematical Intelligencer*, 30(3), 7–17. <https://doi.org/10.1007/BF02985373>
- Mancosu, P. (2018). Explanation in mathematics. In E. N. Zalta (Ed.), *The Stanford encyclopedia of philosophy*. <https://plato.stanford.edu>
- Panza, M., & Molinini, D. (Forthcoming). L'applicabilité des mathématiques. In P. Wagner, M. Panza, F. Poggiolesi, & A. Arana (Eds.), *Précis de philosophie des mathématiques et de la logique*. Editions de la Sorbonne.
- Pincock, C. (2004). A new perspective on the problem of applying mathematics. *Philosophia Mathematica*, 12(2), 135–161.
- Steiner, M. (1998). *The applicability of mathematics as a philosophical problem*. Harvard University Press.

- Urquhart, A. (2008). Mathematics and physics: Strategies of assimilation. In P. Mancosu (Ed.), *The philosophy of mathematical practice* (pp. 588–620). Oxford University Press.
- Wigner, E. P. (1960). The unreasonable effectiveness of mathematics in the natural sciences. *Communications on Pure and Applied Mathematics*, 13(1), 1–14.

ARTIFICIAL INTELLIGENCE

Artificial intelligence (AI) is loosely defined as the capacity of machines that can accomplish tasks that humans would accomplish through thinking. This definition does not say anything about AI achieving this performance in ways similar to how we humans do it, however. The term *artificial intelligence* is used with two meanings. On the one hand, it refers to (artificially) intelligent machines and the ways of making them; in this sense, AI is primarily computer science and engineering. On the other hand, AI is also a transdisciplinary field of studying these machines. AI gurus such as Herbert Simon often emphasized that studying AI involves studying the human mind, and if we get it right, we will understand both AI and the human mind better in the end. Therefore, the field of AI involves various branches of hard sciences and engineering but beyond these also biology, psychology, and philosophy. This entry introduces the types of AI and describes their origins and the ways in which the types differ from each other. Next, the entry pursues the question of what defines intelligence and whether AI machines can be said to *think*. The entry continues with a discussion of the two paradigms of AI research, *strong AI* and *weak AI*. This is followed by an examination of the nature of knowledge and learning and the differences between AI and the human mind—a distinction reinforced in concluding remarks on the importance of this essential consideration in shaping the future of humankind.

Types of AI

The term *artificial intelligence* was coined by the computer scientist and cognitive scientist John McCarthy during a 2-month workshop organized

at Dartmouth College in summer 1956. At the time, only one program (discussed below) qualified for the name, so it was the result of philosophizing about what computers should be capable of. To unpack the concept of AI, it is useful to distinguish between the different types of AI and to delineate facts from beliefs in the process. In this section, the three main types of AI, namely *symbolic reasoning systems*, *symbolic expert systems*, and *artificial neural networks* (ANNs), are examined.

Symbolic Reasoning Systems

The first *thinking machine*—that is, the first AI implementation that worked—was called the *Logic Theorist*, or *Logic Theory Machine*. It was created in 1956 by the Nobel laureate Herbert Simon, his former PhD student Allen Newell, and Cliff Shaw from the RAND Corporation. The purpose of the Logic Theorist was “[. . .] to learn how it is possible to solve difficult problems such as proving mathematical theorems, discovering scientific laws from data, playing chess, or understanding the meaning of English prose” (Newell et al., 1963, p. 109). The Logic Theory Machine gradually evolved into the even more ambitious General Problem Solver, which was supposed to extend the original scope to all areas of human problem-solving. This was the first tentative attempt at what is today called *artificial general intelligence* (AGI).

The departure point was the philosopher Thomas Hobbes’s idea from the 17th century, popularized by AI pioneer Oliver Selfridge (1926–2008), suggesting that all human problem-solving could be represented as a manipulation of symbols. If this was true, Newell and Simon speculated that computers should be able to solve real-world problems, not only arithmetic problems. They tried to achieve this by recording the steps of human problem-solving, using the *thinking aloud* technique: They asked problem solvers to note all the steps of their thinking. Newell and Simon figured that if they repeated the same in many areas of human problem-solving, they should be able to extract the general principles of reasoning, and by applying these principles and generic steps, AI should be able to work across multiple domains. The idea was a radical

departure from the optimization and operational research methods of the time, which started with a detailed model of the complete problem. The performance delivered by the Logic Theorist was astonishing; eventually, it proved 38 of the first 52 theorems in the *Principia Mathematica*, a foundational work in modern symbolic logic by Alfred North Whitehead and Bertrand Russell.

What are the facts and what are the beliefs of symbolic reasoning systems? The first problematic assumption is that all thinking can be represented as symbol manipulation; there is no convincing evidence, so it is just a belief. The second assumption is that there are general principles that apply across all areas of problem-solving; it is difficult to imagine that there are some generic steps that apply from cooking a stew to designing a car or composing music. However, even if this does sound convincing, it is a mere belief. The third assumption is that playing chess or proving theorems, which may arguably involve explicit logical steps, has anything in common with “understanding the meaning of English prose” or anything that involves the achievement of meaning. It is nothing more than a belief. The fact is that the performance of the Logic Theorist was impressive, and since then, similar AI solutions have proved many more theorems and suggested new ones. What has not happened is an AI finding a surprising or significant theorem.

Symbolic Expert Systems

Often called the “father of symbolic expert systems” is Edward Feigenbaum (b. 1936), previously Simon’s PhD student. He abandoned the general principles of reasoning to focus on narrow domains of expertise. He figured that in order to deliver performance on par with a high level of expertise, AI would need to have an internal *model*; a representation of the problem, which would be reasoned. As he began working with scientists, Feigenbaum anchored his project in experimental design in order to make testing possible. Joined by the Nobel laureate Joshua Lederberg, he decided to build a system that would induce molecular topologies from mass spectra, in order to support the Mars probe looking for life, or the precursors of life, on Mars. The project was

named DENDRAL. The input data were reliable, the calculations needed were beyond the usual but not beyond the AI capabilities of the time, and there was a chance to test the findings empirically. The DENDRAL ran for many years, involving an increasing number of experts, eventually covering a larger area than any of the individual experts did. Eventually, the program was performing on par with, and exceeding, top experts.

The DENDRAL was the first *knowledge-based expert system*. A system being *knowledge-based* refers to the internal model, called *knowledge representation*, which is stored in the form of a *knowledge base* (this can be thought of as a cleverly designed database that stores knowledge), while the term *expert system* signifies, not surprisingly, that expert knowledge was represented. The knowledge representation is obtained in the process of *knowledge acquisition*, which is a subset of the overall process of *knowledge engineering*. The facilitator who works with the experts to model their knowledge is called the *knowledge engineer*. As DENDRAL was expanded, the complexity of the knowledge base increased, eventually threatening the stability of the initial LISP (a programming language focused on linked lists) system. Therefore, Feigenbaum’s colleague Bruce G. Buchanan reprogrammed the system, establishing what became the standard for knowledge bases: the “production rules,” or hierarchically organized “if . . . then” rules.

The facts about expert systems are that Feigenbaum was meticulous in selecting a suitable problem for this type of AI: The problem was of the right size, the right level of complexity, experts were available, and there was a possibility of checking the outcome. This is still the crucial feature of expert systems today. In other words, Feigenbaum’s assumptions stood the test of time. Feigenbaum mentioned his belief in AI superintelligence, but this was not used as an assumption for developing expert systems. For many years, expert systems were *isolated*, in the sense that there was no connection and interoperability between knowledge bases; today, knowledge bases can be connected and provide input to each other. The main limit of expert systems today is *brittleness*, meaning that although expert systems deliver high performance in their narrow domains, they are

completely useless even a tiny step beyond their domain's boundary, as there is no representation of nonspecialist knowledge.

Knowledge-based expert systems dominated the AI landscape until the mid-1980s, producing a large number of successful implementations in a wide range of areas from science and manufacturing through to medical treatments.

Artificial Neural Networks

The first ANN was conceptualized in 1943 by Warren McCulloch and Walter Pitts, while the first implementation, the SNARC (Stochastic Neural Analog Reinforcement Calculator), was created by Marvin Minsky and Dean Edmonds as part of their undergraduate work at Princeton University in 1950. The early ANNs were used to calculate mathematical functions; therefore, they should not be regarded as AI but rather as *pre-AI*. The popularity of ANNs rose in the mid-1980s, when computers became sufficiently fast to cope with ANNs of a useful size. Today, ANNs are the most often mentioned forms of AI; they occupy much of the AI landscape, including machine learning.

Any ANN consists of three sets of artificial neurons: There is an input layer, which receives a signal (stimulus); then, there is a hidden layer, which performs a translation or *black-box* operation; and finally, the output layer, which produces a response. The artificial neurons are connected to each other across the layers, resembling synapses in a brain. The response is compared with the expectations, the initial weights are amended, the process is repeated iteratively, and ANN will relatively quickly adjust and produce the desired response. In other words, ANNs learn from a large number of learning examples how to reproduce their statistical frequency. The often-mentioned *deep ANN* and *deep learning* simply mean that there is more than one layer of artificial neurons in the hidden layer.

When it comes to facts and beliefs, it is useful to consider how similar ANN is to the brain. The size of the human brain is estimated at around 80 to 100 billion neurons, each with around 7,000 connections on average, resulting in about 700 to 1,000 trillion (10^{15}) synapses in total. The largest ANN today consists of some 16 million artificial neurons, which roughly corresponds to the size of

a frog's brain. Thus, in terms of synapses, we may be about 8 to 10 orders of magnitude short of the size of the human brain. At the same time, an ANN of 16 million neurons is already so large that the architect cannot grasp it anymore, and it takes a long time to train even using supercomputers. So, if it is even possible to approach the size of the human brain, this will not happen any time soon.

Structurally, ANNs are layered, and the connections usually only go forward (in some networks, there are also within-layer connections). The human brain is a complex network of neurons, so the idea of layered structure may not apply, and there may be circular connections. Functionally, an ANN reflects what was known about biological neurons around the mid-20th century: At any time, neurons are either *firing* or not. Today, it seems that at least a small subset of neurons (perhaps a few million?) display more complex behavior; that is, the strength of the firing impulse seems to matter. Individual neurons also display very different behaviors in this respect, and the strength of firing impulse can vary two orders of magnitude. All this means that structurally and functionally, ANN is rather limited in comparison with the human brain.

However, if ANN architects managed to construct something that resembles an artificial brain in terms of size, structure, and function, there is another belief that must be confronted: that the artificial brain would produce an artificial mind. What is known so far is that brain and mind are somehow connected, and particular thinking processes can be associated with activity in particular brain areas. We still lack an understanding of how the brain and the mind are linked.

Finally, according to recent research besides the brain, other organs, for example, the endocrine system, may be part of our cognitive system. It is not impossible that the mind is fully *embodied*, meaning that the whole body is part of it. So, an artificial brain may or may not produce an artificial mind.

Can AI Think?

Since the very beginning of AI, the big question has been: How do we know if AI should be considered (artificially) intelligent? The first criterion, proposed by Alan Turing in 1950, is still the most popular today. The essence of the Turing

test is that if we communicate with an entity and cannot figure out whether it is a person or a machine, and in fact it is a machine, it means that the machine *thinks*, and it should be considered intelligent. At first sight, this argument may sound convincing; if the machine is not really thinking, we should recognize this. Not everything is, however, as it seems. The first program that passed the Turing test was Joseph Weizenbaum's ELIZA, created in 1966, followed by many other programs, up to today's chatbots. However, these programs passed the test with people who did not know that they were testing. The first experiment that was actually set up as a test was Eugene Goostman, the simulated 13-year-old Ukrainian boy. The legitimacy of this test can be questioned, but, as noted by Stuart Russell and Peter Norvig, it fooled 33% of the untrained amateur judges. So, the judges might not be right and 33% is quite low. In addition, the setup certainly worked in favor of the outcome: having someone who is 13, from a country about which the testers do not know much, speaking English as a foreign language. As Russell and Norvig speculated, perhaps the Turing test is really a test of human gullibility.

It is possible, however, that an AI will at some point legitimately pass the Turing test—but there are further problems. Turing modeled his test after the *imitation game*. In this game, Person *A* pretends to be Person *B*. If *A* has learned a lot about *B*, the testers may believe that *A* is *B*. However, this does not mean that *A* actually became *B*. The programs that passed the Turing test were designed with the purpose to pass the test, not with the purpose to think.

An excellent explanation of the Turing test is the *Chinese room argument*, proposed by the philosopher John Searle. The essence of the argument is the following: A person who does not speak Chinese is in a room full of rule books, receiving messages in the form of Chinese symbols and looking up responses to those messages in the rule books. If the rule books are good, those reading the responses outside the room may believe that the person responding speaks Chinese. Similarly, passing the Turing test does not prove that the computer thinks, only that the program is good.

Of course, none of these instances prove that machines cannot think; they only point out the

lack of conclusive proof. So, we are left to our respective beliefs.

AI Paradigms

The distinction between *strong AI* and *weak AI*, as the two paradigms of AI research, was introduced by Searle, the author of the Chinese Room Argument. The strong AI paradigm postulates the idea of the thinking machine—of AI with a mind of its own, at least—as a desirable and achievable outcome. Most followers of strong AI do not believe that AI can currently actually think, but they do believe that it is only a matter of time until it will. Typically, this is fueled by the beliefs described in the previous and following sections of this entry. Marvin Minsky, a great proponent of strong AI, suggested that any endeavors in the field of AI, other than trying to create AGI, are meaningless.

In contrast, those who follow the weak AI paradigm see AI as a technology that can help produce useful and powerful tools. Many of them do not believe in the possibility of AGI, and none of them work on the AGI project. They argue that we cannot produce an artificial mind, as it would require the researchers to obtain a complete understanding of the mind, which they believe to be impossible. There is a very important consequence of this belief on what and how the followers of weak AI try to achieve in their work. They do not try to replicate the human thinking process; instead, they try to find a sensible way of achieving the same or similar performance. A popular metaphor used to describe this is the relationship between the horse and the car. When the original creators dreamed up the first automobile, they did not attempt to replicate the muscles, joints, and metabolism of a horse; they used an internal combustion engine instead, and put it on wheels. The car can be an excellent replacement for some aspects of the horse; for instance, it can take passengers from one place to another (more passengers than a horse, more conveniently and usually faster). In other areas, however, such as playing polo or enjoying horseback riding, the car cannot deliver. Similarly, what weak AI researchers produce only works in a narrow area and in very different ways from how humans would deal with the task.

Curiously, although their beliefs are fundamentally different and they focus on different types of work in the area of AI, the followers of the two paradigms do agree about one thing: that what is missing to achieve a thinking machine is common sense. As all human beings have common sense, it is sometimes considered to be something simple—but it is not. According to Minsky (1988, p. 22), common sense consists of “hard-earned practical ideas—of multitudes of life-learned rules and exceptions, dispositions and tendencies, balances and checks.” It only appears simple because we do not recall acquiring it.

AI and the Human Mind

There are several areas that people think of as particularly human, such as knowledge, learning, and creativity. It is useful to look at the differences between the human mind and AI along these dimensions that can be conclusively demonstrated.

The difference that is the easiest to identify along the knowledge dimension is that systems that rely on knowledge acquisition are necessarily limited to explicit knowledge. According to the Hungarian–British polymath Michael Polányi, all human knowledge is either tacit or rooted in that which is tacit, as all explicit knowledge relies on being tacitly understood and applied. The problem of common sense showcased in the previous section can be explained similarly: Common sense is predominantly tacit, and all human knowledge incorporates common sense. *Intuition*, which has been deemed valuable in relation to peak human achievements (such as scientific discovery, artistic creativity, and managerial decisions), also belongs to the tacit realm. Humans also know *harmony and beauty*, and this beauty usually drives creative achievements, deep thoughts, and simply provides joy. For instance, a human can admire the beauty of a rainbow despite knowing that it is the result of the prism effect. People also use *analogical thinking*, which refers to mental models analogical to reality. An extraordinary example of this is the Serbian–American inventor Nikola Tesla (1856–1943) designing his machines entirely in his mind, without using a blueprint. Finally, *seeing the essence* is an achievement unique to the high levels of human expertise. It consists of two parts: the *big picture* and the detail. Those who achieved

a high level of expertise are known to see the *big picture*. However, they can also see any of the details, as well as the relationship between the details as well as the relationship of the details to the big picture, and they can rapidly switch from one to the other.

Besides the many aspects of knowledge that do not seem to be easy to replicate or emulate with AI, knowledge also relates to other aspects of our being. Humans also have *feelings*, which are manifestations of biological drives such as hunger, fear, longing for the company of other people, and so on. While humans can temporarily override their feelings, using their knowledge—for example, we do not need to eat right away when we start feeling hungry—*emotions*, in turn, can override knowledge. Our important decisions, such as whom to marry, where to live, or what organization to work for, have significant emotional components. Minsky goes so far as to suggest that knowledge is impossible without emotions. Humans also have *values* embedded in their knowledge. For instance, sometimes they do not do what matches their well understood preferences but what they think is the right thing to do.

The situation with learning is similar to the situation with knowledge. Reinforcement learning, which is how ANN learns, is only a tiny part of human learning. There was a time when psychologists thought that all human learning is a variant of reinforcement learning, but psychology has come a long way over the past 70 years. The significance of talent has been discovered, the significance of which is that people learn faster in areas in which they are talented, and learning in these areas feels mostly like playing. AI learns equally in all areas. People get inspired and interested in various things, and inspiration and interest can make them invest extra efforts in learning what they like. Furthermore, those who achieved the highest level of expertise all went through some form of master–apprentice relationship. All people have some such experience; for instance, this is how we learn our native language. This immensely complex form of learning builds on a deep personal connection between master and apprentice, and our understanding of how it works is very limited. What we do know, however, is that the master–apprentice relationship seems to be the only way of transferring tacit knowledge.

Finally, there is a question if AI can be creative or not. It certainly outperformed top human experts in areas where creativity is considered important. Examples are Deep Blue winning against chess grandmaster Garry Kasparov in 1997, and AlphaGo defeating Lee Sedol in 2016 (as well as all the other Go grandmasters shortly after that). Human creativity is defined as the production of a new and useful idea—AI has definitely produced new and useful things, but it can be argued that it has never produced an idea. However, AI can certainly support human creativity, producing patterns from large amounts of data as well as helping us transcend the boundaries of our knowledge traditions.

Concluding Remarks: How to Think and Talk About AI

This entry has described what AI can do and what its limitations are; the entry has also delineated the facts and beliefs regarding current performance and future promises. AI comprises an exceptionally powerful set of technologies and may well be the most expensive human enterprise ever. If we use it well, it can greatly benefit the whole of humankind.

What can be done to get things right about AI? Perhaps it would be a good start to stop describing AI in human terms: *Human–AI collaboration, thinking, learning, and making decisions* are all terms that attribute intentionality to AI—something that it most definitely does not currently have, and possibly never will. Even when we talk about authentic AI, we usually think of a better fake human. Authentic AI, however, is not human-like; it is AI-like. For a human future in which AI plays a great role to the benefit of humankind, we need authentic humans as well as authentic AI.

Viktor Dörfler

See also Analysis; Cognitive Science; Intuition; Knowledge; Philosophy of Mind; Rationality; Reasoning

Further Readings

- Davenport, T. H. (2018). *The AI advantage: How to put the artificial intelligence revolution to work*. MIT Press.
- Dörfler, V. (2018). *Big data, artificial intelligence and the human mind*. Paper presented at the TEDx University

- of Strathclyde, Glasgow, UK. <https://youtu.be/KAXfo-cZ80Y>
- Dörfler, V. (2020). Artificial Intelligence. In M. A. Runco & S. R. Pritzker (Eds.), *Encyclopedia of creativity* (3rd ed., Vol. 1, pp. 57–64). Academic Press.
- Dörfler, V. (2022). *What every CEO should know about AI*. Cambridge University Press.
- Dreyfus, H. L., & Dreyfus, S. E. (2000). *Mind over machine: The power of human intuition and expertise in the era of the computer*. The Free Press (Original work published 1986).
- Feigenbaum, E. A. (1992). *A personal view of expert systems: Looking back and looking ahead*. Stanford University. <https://purl.stanford.edu/gr891tb5766>
- McCorduck, P. (2004). *Machines who think: A personal inquiry into the history and prospects of artificial intelligence* (2nd ed.). A K Peters.
- McGilchrist, I. (2019). *The master and his emissary: The divided brain and the making of the western world* (2nd ed.). Yale University Press.
- Minsky, M. L. (1988). *The society of mind*. Simon & Schuster.
- Newell, A., Shaw, J. C., & Simon, H. A. (1963). Empirical explorations with the logic theory machine: A case study in heuristics. In E. A. Feigenbaum & J. Feldman (Eds.), *Computers and thought* (pp. 109–133). McGraw-Hill.
- Penrose, R. (2016). *The emperor's new mind: Concerning computers, minds and the laws of physics*. Oxford University Press (Original work published 1989).
- Polányi, M. (2002). *Personal knowledge: Towards a post-critical philosophy*. Routledge (Original work published 1962).
- Russell, S., & Norvig, P. (2020). *Artificial intelligence: A modern approach* (4th ed.). Pearson Education.
- Searle, J. R. (2004). *Mind: A brief introduction*. Oxford University Press.
- Simon, H. A. (1996). *The sciences of the artificial* (3rd ed.). MIT Press.
- Winograd, T., & Flores, F. (1987). *Understanding computers and cognition: A new foundation for design* (2nd ed.). Addison-Wesley.

ASSUMPTION

Our contemporary word *assumption* is derived from the Latin word referring to the act of being taken up or received. In that context, the Catholic Church used the word to articulate the Assumption

of the Virgin Mary—that is, the official dogma that she was taken up bodily into Heaven. That use remains today, but beginning in the 13th century the word slowly took on the broader meaning of any idea, premise, or axiom that is *taken up* or taken for granted. With this latter meaning in mind, in what follows, a definition of assumption is offered and explained point by point, giving examples along the way as the term relates to both theory and human life. From there, a few additional examples are provided of the importance and place of assumptions in theory and theory building, spanning not only the physical sciences and mathematics but also the social sciences. The entry concludes with a brief summation of the place and importance of assumptions in theories.

What Is an Assumption?

Assumptions are beliefs of a particular kind—namely, assumptions are propositions that persons (and in the aggregate, human groups or even entire societies) regard to be true, whether consciously or tacitly, and which form the basis of—rather than being conclusively inferred from—processes of reasoning and lived experience. This definition requires some unpacking. First, assumptions are propositional, but oftentimes this is the case merely in the sense that, in principle, they could logically be articulated as propositional statements (unlike, for instance, some kinds of emotional states). In other words, assumptions are propositional not in the sense that they are necessarily or even usually propounded and spelled out in explicit propositional form. But assumptions are propositional in the sense that they make claims regarding what is true about reality and those understandings—even when left unstated—can be either true or false.

Secondly, assumptions are not free-floating ideas but are believed by persons and often also by groups, subcultures, or societies. It is not hard to grasp that persons hold fast to assumptions. But when persons gather, interact, or even share some common social identity from a distance, their association is likely to give rise to (or be premised upon) taken-for-granted ideas, values, and axioms. The first-generation college student group at the state university likely holds certain assumptions—about class mobility, equality, and

the purpose of higher education—which take intentional work to see and to understand. It is also possible for the vast majority of a society to hold to an assumption that is rejected by a subcultural subset within it. For example, the popular assumption in Western modernity that what is morally right and good is ultimately up to each individual to decide is rejected by many conservative religious thinkers and some philosophers.

Third, assumptions can function anywhere along a spectrum from conscious to routinized to subconscious, being explicit or left implicit. In some cases, assumptions are stated up front and explicitly recognized as providing the grounds for some subsequent argument, observation, or experience. This is the case, for instance, in many philosophical arguments (especially those in analytic philosophy using formal logic) as well as in the mathematical proofs, one undertakes in algebra or geometry classrooms. In this sense, assumptions are rightly understood as taken-for-granted premises, the logical givens that set the stage for what comes next in the course of making a case or building an argument. In other cases, assumptions are known but nevertheless left unexpressed. For example, an evangelical Christian might try to organize her life and beliefs around the idea that the Bible alone is God's authoritative Word to humans. Such a baseline assumption would presumably have a powerful guiding effect on her life and beliefs even if most days it does not come up in conversation, but if asked she would be readily able to express it.

At still other times, assumptions operate on a level that is wholly unrecognized and implicit. Persons can hold assumptions without even knowing that they hold them, and persons can be powerfully influenced by those assumptions without recognizing it. Assumptions—as a particular kind of belief—need not be reflected upon, consciously adopted, or able to be stated in order to be genuinely believed. This does not mean, though, that such assumptions must always stay at the prediscursive, implicit level of consciousness. It is often possible for another person to point out and articulate a previously unrecognized assumption, at which time that assumption might move from unrecognized to recognized, from implicit to explicit. It is even possible for a person to identify her own previously unrecognized assumptions and change them if necessary.

Sometimes this is difficult, requiring study, dialogue, and critical reflection. Other times, this does not take much. Persons often do not realize they had been holding some assumption until it turns out to be false: You return home from a weeklong vacation only to discover that you had wrongly assumed that your spouse had locked the front door. You arrive at your favorite coffee shop to meet a friend, not having considered that she meant the one on the other side of town.

Fourth, assumptions are not (and cannot be) conclusively inferred or verified from reason or experience. Against the tradition of rationalism, taking assumptions seriously in life and in academic theorizing requires the recognition that reason, rationality, and logic are never the only factors at play. Likewise, against the tradition of empiricism, seriously accounting for the role of assumptions in life and theorizing necessitates moving beyond a strict focus only on what can be observed from experience and the five senses. Instead, assumptions are precommitments and premises that come before either reason or empirical experience. In the language of epistemology (the philosophy of knowledge), an assumption can be *justified* (i.e., a person is not always being cognitively reckless or intellectually sloppy for believing it) without being *justifiable* (i.e., without being able to be verified through argumentation and evidence). This does not mean that assumptions cannot be supported or made compelling through reasoning and experience. If an assumption meshes well with one's lived experience and helps to make the world intelligible, then there is good reason to hold to it. It is justified. But it cannot be conclusively established or deduced with indubitable certainty through the use of reason or observation. It is not justifiable.

Fifthly, basic assumptions form the grounds for processes of reasoning—ranging from particular instances of philosophical argumentation and mathematical proofs to more general approaches to (or convictions about) what knowledge is and how humans can go about acquiring it. This is most easily seen in simple logical syllogisms, in which a major premise and a minor premise are used to deduce a conclusion. But even in more complex belief structures and systems of reasoning, assumptions can serve the role of basic (and powerfully life-guiding) beliefs. That is, certain

kinds of beliefs—such as that a divine being exists, that one's memories really do refer to actual past happenings, or that natural science is the only rational way to reach conclusions—can *sit at the bottom*, so to speak, of one's system of beliefs or worldview. Such basic beliefs are not inferred from, or held on the grounds of, other beliefs that the knower regard to be true. And different persons (and human groups) can and do hold to differing beliefs in this basic way. This pluralistic nature of basic starting points has led many philosophers of knowledge away from foundationalism (i.e., the theory that all humans share some common, incorrigible foundation for knowledge) to what is known as *post-foundationalism*. This is also why it might seem like a life-altering paradigm shift when one's basic beliefs and assumptions, for some reason, are revised.

Sixth, assumptions and premises are crucial not only for theory and theorizing but also for human experience and life in the world more broadly. The propositional content of assumptions can range from the mundane (“Oh, I assumed you were going to buy the milk”) to big questions in the realms of epistemology, metaphysics, and ethics. And to the extent that abstract ethical, epistemological, and metaphysical issues and ways of thinking influence the ways persons live, even seemingly distant assumptions turn out to have serious consequences on actual experience. On this point, some philosophers and social scientists emphasize the principle of *practical adequacy*, which suggests that when the components of human knowledge systems (including one's assumptions) *rub up against* practical activity in the world in a way that makes the world and life unintelligible or unlivable, there is good reason to suspect that some crucial part of that knowledge system is problematic (i.e., false) and thus may need to be reconsidered and revised.

Examples of Assumptions

Assumptions play a central role in theories and theory building in several ways. In the social sciences, for example, assumptions about the nature (or lack thereof) of human beings are multiple and lead theorists in different directions. Some theorists hold a social collectivist view that claims human beings are ultimately a product of the

cultural environments in which they are embedded, such that they are primarily concerned with living out their society's norms and values. Other theorists assume a fundamentally economic view of humans as self-interested rational actors, seeing humans as driven by cost-benefit calculations in order to maximize their utility. Other theorists, influenced by postmodernism and phenomenology, view human identity and selfhood as fragmented, transient, and variably constructed across different situations—lacking any essential nature. Still other theorists adopt a sociobiological and evolutionary view of humanity, in which all of human social life can be explained by reference to *survival of the fittest* and the functioning of genetic material. And some theorists hold to an Aristotelian view of human beings as persons with an essential shared human nature and a moral direction toward which to strive for their objective flourishing. Which of these models of human nature is taken for granted clearly would have a great impact on a social scientist's explanations and theories.

Another example can be drawn from the history of geometry. In the third century B.C.E., the Greek mathematician Euclid systematized geometry based on five axioms, from which more complex theorems could be deduced. These five axioms involved the relationships between lines, points, and angles (e.g., between any two points there is only one line segment with those points as end points). For more than 2,000 years, this axiomatic system was the only version of geometry. But over the centuries, several thinkers questioned whether Euclid's fifth axiom (namely, that for any line and a given point off that line, there is only one line that runs through that point and parallel to the first line) was self-evidently true. Several attempts to prove the fifth axiom based on the first four ended in failure. Beginning in the early 1800s, mathematicians started to imagine what alternative geometric systems would look like if Euclid's fifth axiom—known as the *parallel postulate*—could be false. The result was a variety of non-Euclidean geometries, in which the spatial grid on which geometry is done is not a boxy, straight *Euclidean space*, but instead is intrinsically curved in various ways. This questioning of Euclid's fifth assumption and the recognition that spatial grids can be curved helped to shift the way physicists think about space and

set the stage for Albert Einstein's theory of general relativity and gravitational force.

Even the scientific method operates on the basis of a handful of assumptions that cannot be verified using the scientific method. For example, science proceeds on the presupposition that there exists a real world about which scientists are trying to gain knowledge and that this world has a structure and nature independent of our theories about it. Scientific inquiry and discovery also makes sense only if scientists believe that aspects of the physical world (such as the laws of physics or the basic features of biology) will work very similarly tomorrow as they have worked in the recent past. Likewise, the scientific method assumes that human cognitive and sensory faculties are fairly reliable tools for gathering and interpreting data. Even more basically, science proceeds under the supposition that some beliefs and theories are closer to the truth than others; all is not relative. Science even makes moral assumptions—such as that theories should be tested fairly, that findings should be reported honestly, and that it is better to believe what is true than what is false. As assumptions, such beliefs are not necessarily unjustified or wrong, but they are nevertheless not definitively verifiable by logical proofs, arguments, or empirical observations.

Conclusion

It is not necessary—and probably not possible—for every unverified assumption to be articulated in theory and research. But clearly, assumptions are both important and unavoidable. While some scholars may feel the urge to hide, ignore, or deny the premises and assumptions that underpin their work, it is best as much as possible to be aware of the assumptions that one is making. And beyond just reflecting upon and recognizing one's assumptions, it is important as well to be willing to question and possibly to change them when the data and evidence call for it. Even the most fair-minded scholarship does not operate with a God's-eye view or with every idea or belief established with absolute certainty. But with sustained and critical attention to the assumptions undergirding scholarship, development of theory can proceed on a steady footing.

Brad Vermurlen

See also Belief Revision; Epistemology; Warrant; Worldview

Further Readings

- Alston, W. P. (2005). *Beyond "Justification": Dimensions of epistemic evaluation*. Cornell University Press.
- Llewelyn, J. E. (1962). Presuppositions, assumptions, and presumptions. *Theoria*, 28(2), 158–172.
- Schlefer, J. (2012). *The assumptions economists make*. Belknap Press.
- Smith, C. (2003). *Moral, believing animals: Human personhood and culture*. Oxford University Press.
- Ullman-Margalit, E. (1983). On presumption. *Journal of Philosophy*, 80(3), 143–163.

ASTRONOMY

See Cosmology

AUTHORITY

The concept of authority is fundamentally related to the concepts of hierarchy, power, and obedience. In modern research on authority, the concept has been articulated and operationalized in a number of different ways. Scientific discourse is discussed in terms of its influence and paradigmatic power, intrinsically linked to its authoritativeness. Political scientists refer to systems of authority and how they relate the role of the state. Psychologists refer to the relationship between authority and obedient behavior.

The word *authority* is etymologically rooted in Latin and is connected with the words *auctoritas* and *auctor*, which reference “one who originates or promotes ideas.” In Old French, the term *autorité* is itself related to the Old English term for *author*, again referencing “one who originates.” These roots are reflected in the general concept of authority as “one who originates or issues command or orders,” ostensibly to be obeyed by those subservient to the person, their social role, and/or the orders. Conceptually speaking, these dynamics naturally invokes reference to the constructs of hierarchy and the power differentials that underlie

and support a hierarchical social system. Philippe Aghion and Jean Tirole note the distinction between *formal* and *real* authority in organizations, the former being ascribed to an office who holds power by virtue of the position (e.g., a chief executive officer, or CEO, who may not actually make real decisions impacting organizational goals) and the latter having the ability to make decisions on the basis of authority granted by the superior.

This entry considers the varying conceptualization of authority, and how they have been reified in scientific, political, religious, and psychological narratives. This list is by no means exhaustive but is meant to synopsise some of the most prominent discourses that have been developed around explicating authority and how it structures human behavior.

Scientific Authority

Scientific discourse is powerful and influential to the extent that it is considered authoritative. Thomas Kuhn described the manner in which scientific theories come to dominate others and eventually become widely accepted as lenses through which problems are both conceptualized and solved. In his influential book *The Structure of Scientific Revolutions*, published in 1970, Kuhn describes the social process whereby scientific theories become authoritative, a point to which he ascribes the term *paradigm*. Scientific paradigms do change over time, largely due to the presentation of anomalies that cannot adequately be addressed or resolved within the context of the dominant paradigm of the time. When this happens—and Kuhn suggests that it is ultimately inevitable—a paradigm shift occurs whereby novel, unanticipated solutions to the anomalies arise, and achieve acceptance in the wider scientific community. Over time, the new paradigm comes to be accepted as authoritative among the scientific community and society in general.

Political Authority

Political authority is a concept that is intrinsically linked with the concepts of power and legitimacy. Power, in its most abstract conceptualization, typically entails the ability to influence others behavior beyond what they would ordinarily do, and

legitimacy is most often conceptualized as a form of government which citizens believe is appropriate for that society. C. W. Cassinelli noted that legitimacy and political authority almost always involve an element of coercion, but can never involve force or the explicit threat of it; to do so is to undermine the extent to which one is truly considered a political authority.

Max Weber (1864–1920) developed a typology of political authority in his treatise “The Three Types of Legitimate Rule”: traditional, legal-rational, and charismatic. Weber’s three types of authority represent ideals, and thus do not necessarily correspond precisely with any given socio-political entity. In addition, authority can evolve from the lowest form into one of the other, more advanced, forms. *Charismatic authority* is described as a social arrangement wherein leaders derive their power on the basis of their possession of some type of extraordinary attribute or supernatural power. Those subject to this type of rule must inherently believe in the power granted the leader. A contemporary example of a state predicated upon charismatic authority is North Korea under the totalitarian rule of the Kim family. The Kim family began their reign with Kim Il Sung, a leader whose persona was so elaborated by means of state-controlled propaganda, he continued to serve for years after his death as the world’s only posthumous president.

Traditional authority arises from longstanding custom and is the basis of authority for monarchies. The death of a charismatic leader can result in the formation of a system whose authority derives from tradition. This type of authority is oriented toward maintenance of the status quo ante and its inherent inequalities. Beyond the domain of nation-states and their form of political organization, an example of this type of authority can be found in patriarchal societies, wherein the inherent inequality between men and women has been established not on the basis of charisma (or rational principles) but solely on the basis of longstanding custom.

The *legal-rational system of authority* is organized such that citizens are subservient to a legal structure that has been rationally constituted, rather than to a particular person, as is the case within a system predicated upon a charismatic or traditional form of authority. That legal structure

can take the form of a constitution, a static form of bureaucratic organization, or rationally derived codified principles that establish the legality, responsibilities, and conditions of rule. Typically, under this form of authority, citizens are deferential to the office wherein power is vested, and not to the particular officeholder who happens to fulfill that social role. As an example, 21st-century Japanese, British, or Canadian political authority would be characterized as legal-rational in organization.

Religious Authority

Mark Chaves defines religious authority as “a social structure that seeks to enforce its order through the legitimate control of some supernatural component” (1994, p. 750); where authority withholds access to something vis-à-vis the legitimacy of the supernatural, that authority is religious. The goods that religious authority might control vary between religious traditions (e.g. eternal life, an end to suffering, or prosperity). Authoritarianism has been linked with religiosity and is a social attitude that refers to individuals who hold conservative principles and are generally accepting of established authorities. Authoritarians and deeply religious people tend to share common values, preferring the status quo over change, and employing a noninterpersonal morality that refers to external sources of authority (e.g., group norms).

Psychological Authority

The psychological literature has considered the nature of authority as it manifests itself in interpersonal relationships (such as between parent and child) and how authority is related to obedient behavior. The scope of discourse surrounding parental authority generally concerns parents’ ability to utilize one of two types of control: psychological control and behavioral control. Psychological control involves any attempt by parents to encroach upon a child’s psychological and emotional development. It typically takes the form of manipulative tactics in order to achieve its goal. Behavioral control refers to parental attempts at managing a child’s behavior, such as correcting a child’s inappropriate behavior with punishment.

Psychological control is most potent in young people before and during adolescence and is sometimes thought to be a parenting style as opposed to a parenting practice. This type of control does not necessarily involve parental pressure to feel and think in ways dictated by parents; it can merely embody parental pressure aimed at making the child behave in accordance with parental expectations. Tactics like guilt induction help parents maintain legitimate parental authority over issues of concern, like child friendship affiliations. Parental control can sometimes intrude upon a child's psychological development. Where it does, empirical evidence suggests that the negative consequences of this control can be predictive of later mental health issues and behavioral disorders. Where it concerns behavioral control, research on parental authority suggests that it is not counterproductive unless it impairs the bond between parent and child; harsh or erratic levels of punishment are predictive of later involvement in activities such as delinquent behavior.

Parents are encouraged to balance appropriate control over their children's behavior with developmentally appropriate attempts to give them autonomy as they grow older. Younger children are more accepting of parental authority and its legitimacy; however, some have found that as children aged, their views on parental authority become less positive. Where it concerns the scope of parental authority, there appear to be clear boundaries with regard to what constitutes legitimate parental authority and what should remain under the jurisdiction of children.

Stanley Milgram and *Obedience to Authority*

One of the most controversial and thought-provoking research projects aimed at understanding the nature of authority was psychologist Stanley Milgram's work at Yale University concerning obedience to authority figures. His research, first described in the *Journal of Abnormal and Social Psychology*, was published contemporaneously with Hannah Arendt's coverage in 1961 of the Adolf Eichmann trial for *The New Yorker* magazine, the hugely consequential trial of a former German bureaucrat instrumental in supervising, organizing, and facilitating the systematic extermination of

European Jews, or *Jewish Affairs*, as it was euphemistically put in German bureaucratese. Fifteen years after the close of the war, Eichmann had been discovered living in Argentina by the Simon Wiesenthal organization; he was subsequently abducted by Mossad, the Israeli intelligence agency and smuggled back to Israel on a commercial flight to stand trial as personally accountable for his role in the murder of 6 million Jews.

Milgram was explicit about his interest in exploring the social dynamics that had facilitated the implementation of the Final Solution to the Jewish Question, the event now formally referred to as *the Holocaust*, and intended to compare levels of compliance in Germany with those in the United States. Aside from inspiring an onslaught of critique from research ethicists, his 1974 book *Obedience to Authority: An Experimental View* exposed a social dynamic whereby free will can be truncated, overridden, in the name of obedience to those who are perceived to constitute legitimate authorities.

The experimental design saw to it that experimental subjects administered an incrementally more intense series of electric punishment shocks to an ostensible *subject* (in reality a confederate of the experimenter) under the guise of its being a test of negative punishment's effect on learning capability. The coercive prompting of a confederate *researcher*, playing the role of a legitimate authority figure, insisted at appropriate times that subjects continue with the experiment despite any protests they might have to delivering increasingly strong electrical shocks to the victim, the bogus *subject*. Milgram performed many permutations of the relationship between the various actors involved in the experiment—altering in varying degrees the proximity of the victim, confederate researcher, and the experimental subject in order to test whether social distancing had an impact on a subject's obedience to orders, which was generally found to be the case.

Replications of the shock experiments were canceled in Hamburg, Germany, when Milgram found levels of compliance in the United States (the control group) high enough to rule out meaningful comparison with what he originally conjectured would be a more obedient German population. To the lament of the predictions of the psychiatrists Milgram had surveyed beforehand, roughly two-thirds (66%) of the subjects in Milgram's experiments were fully compliant in

harming (and ostensibly killing) the confederate victim. The dynamic that Milgram exposed was one whereby personally injurious behavior need not be related to any moral, psychological, or biological factors impacting the experimental subject. Milgram's overall findings are compatible with the sincerity of the defense cliché espoused by those tried for offenses against humanity at Nuremberg and elsewhere: that they were merely *following the orders* of their superiors, presumably because they did not have the grounds to question the legitimacy of that authority, or the commands issued forth.

Douglas J. Dallier and Olivia M. Lewis

See also Evidence; Fact Versus Theory; Warrant

Further Readings

- Aghion, P., & Tirole, J. (1997). Formal and real authority in organizations. *Journal of Political Economy*, 105(1), 1–29.
- Asquith, P., & Smetana, J. G. (1994). Adolescents' and parents' conceptions of parental authority and personal autonomy. *Child Development*, 65(4), 1147–1162.
- Barber, B. K. (1996). Parental psychological control: Revisiting a neglected construct. *Child Development*, 67(6), 3296–3319.
- Cassinelli, C. W. (1961). Political authority: Its exercise and possession. *Political Research Quarterly*, 14(3), 635–646.
- Chaves, M. (1994). Secularization as declining religious authority. *Social Forces*, 72(3), 749–774.
- Kuhn, T. S. (1970). *The structure of scientific revolutions*. University of Chicago Press.
- Milgram, (1969). *Obedience to authority; An experimental view*. Harper & Row.
- Weber, M. (1958). The three types of legitimate rule. *Berkeley Publications in Society and Institutions*, 4(1), 1–11.

AXIOM SCHEMA

An *axiom schema* (plural: *schemata* or *schemas*, from the Greek *skhēma*, meaning *form* or *figure*) is a *template* for an axiom, accompanied by a *rule* that tells us how to instantiate the template with

specific expressions in some (formal) language to produce an axiom. Thus, in developing an axiomatic theory in logic, mathematics, or science, one uses axiom schemas to specify a collection of axioms representing the basic or fundamental assumptions of the theory, from which one derives all other implications of the theory (i.e., its *theorems*). However, there will typically be *infinitely* many such axioms and so some method is needed to specify all of them at once by finite means. It is primarily because of this that we require the templates for axioms known as *axiom schemas*.

It is worth noting that *axiom schemas* are just one instance of the slightly broader notion of a schema, used in both logic and mathematics. The broader notion includes templates both for the axioms of a theory and so-called *rules of inference*, that is, rules for deducing consequences from the theory's axioms or from previously established theorems or previously made assumptions. Because there will also be an infinite number of ways of applying a given rule of inference, a schema for representing this rule by finite means is also needed. This entry discusses examples of axiom schemas in both logic and mathematics.

Axiom Schemas in Logic

Axiomatic systems in logic, often called *formal systems* or *logical calculi*, allow us to determine which statements in a formal language are logical validities, either because they are axioms themselves or because they are logical consequences of axioms. For example, in a *Hilbert system* for propositional logic, an approach that emphasizes axioms and minimizes the role of deductive inference rules, the following expressions are examples of axiom schemas:

$$P1: (\phi \rightarrow \phi)$$

$$P2: (\phi \rightarrow (\psi \rightarrow \phi))$$

$$P3: ((\phi \rightarrow (\psi \rightarrow \chi)) \rightarrow ((\phi \rightarrow \psi) \rightarrow (\phi \rightarrow \chi)))$$

$$P4: ((\sim(\phi) \rightarrow \sim(\psi)) \rightarrow (\psi \rightarrow \phi))$$

In each of these four axiom schemas, we have expressions representing logical connectives; here

the arrow, “ ρ ,” represents material implication and the tilde “ $\sim$ ” represents negation. Other systems might also involve symbols for disjunction (“ $\vee$ ”), conjunction (“ $\wedge$ ”), and the biconditional (“ $\Leftrightarrow$ ”). Technically, the left and right parentheses are also part of this logical vocabulary, necessary for indicating and disambiguating the scope of the logical constants. We also have the expressions “ ϕ ,” “ ψ ,” and “ χ ” representing arbitrary well-formed formulae in the formal language in which the system is expressed. These latter expressions are sometimes called *metalinguistic variables* as they are not themselves expressions in the formal language of the system but are expressions from the *metalanguage*, the language we use to talk about and study the formal language and properties of the system. Thus, in contrast to the metalanguage, the language of the system itself is called the *object language*, and, in the context of propositional logic, the metalinguistic variables range over the class of well-formed formulae (or *wffs*) in the object language. Most commonly either Greek letters or uppercase Latin letters ($A, B, C, \dots$) are used to denote metalinguistic variables.

Each of P1 through P4 is thus a template in the sense of being a string of symbols that outlines and represents the shape or form that an axiom in the Hilbert system of propositional logic can take. However, these axiom schemas also implicitly involve a *rule* indicating how they are to be instantiated or *filled in* for the sake of producing a genuine axiom expressed in the language of the system itself. That rule is implicit in the range of values that the metalinguistic variables, “ ϕ ,” “ ψ ,” and “ χ ” can take, which is here stipulated to be the class of well-formed formulae of the formal language of propositional logic. To show what these instances might look like we can first briefly sketch a formal language for the propositional logic embodied in this Hilbert System:

Let L_{PROP} be the language of propositional logic, defined as the following set of wffs:

- i. $P_i \in L_{PROP}$, for every $i \in \mathbb{N}$
- ii. If $\phi \in L_{PROP}$ then $\sim(\phi) \in L_{PROP}$
- iii. If $\phi, \psi \in L_{PROP}$ then $(\phi \rightarrow \psi) \in L_{PROP}$

- iv. Only expressions formed by applications of (i)–(iii) are wffs of L_{PROP} .

Thus, applying these formation rules, the following strings of symbols are all examples of well-formed formulae in L_{PROP} :

- a. $P_1 \rightarrow (P_4 \rightarrow P_{23})$
- b. $\sim (P_2 \rightarrow P_{217}) \rightarrow (P_{1300} \rightarrow (P_{14723} \rightarrow \sim (P_0)))$
- c. $P_5 \rightarrow ((P_7 \rightarrow (\sim P_9 \rightarrow P_{11})))$

Procedurally, having first specified a formal language for propositional logic so that we now know what counts as a wff, the axiom schemas P1 through P4 are then introduced to specify which specific wffs in the language count as axioms of the deductive system. The following are therefore axioms of the system as they are instances of the axiom schemas P1 through P4, respectively:

$$\begin{aligned} &(P_7 \rightarrow P_7) \\ &(P_7 \rightarrow (P_{10} \rightarrow P_7)) \\ &(P_2 \rightarrow (P_{111} \rightarrow P_{6455})) \rightarrow \\ &((P_2 \rightarrow P_{111}) \rightarrow (P_2 \rightarrow P_{6455})) \\ &((\sim P_5 \rightarrow \sim P_{108044}) \rightarrow (P_{108044} \rightarrow P_5)) \end{aligned}$$

The following more complicated formulae are also axioms and instances of P1 through P4, respectively:

$$\begin{aligned} &(P_7 \rightarrow (P_{42} \rightarrow (P_{500} \rightarrow P_{99}))) \rightarrow \\ &P_7 \rightarrow (P_{42} \rightarrow (P_{500} \rightarrow P_{99})) \\ &(\sim (P_{80} \rightarrow P_7) \rightarrow ((P_{10} \rightarrow P_{778}) \rightarrow \\ &\sim (P_{80} \rightarrow P_7))) \\ &(P_2 \rightarrow (\sim (P_{111}) \rightarrow \sim (P_6 \rightarrow P_{89}))) \rightarrow \\ &((P_2 \rightarrow \sim (P_{111})) \rightarrow (P_2 \rightarrow \sim (P_6 \rightarrow P_{89}))) \\ &((\sim (P_7 \rightarrow (P_{10} \rightarrow P_7)) \rightarrow \sim (P_{10} \rightarrow P_{778})) \rightarrow \\ &((P_{10} \rightarrow P_{778}) \rightarrow (P_7 \rightarrow (P_{10} \rightarrow P_7)))) \end{aligned}$$

The axioms can thus become almost arbitrarily complex. In a Hilbert system for propositional logic, the axiom schemas are typically also supplemented with a schema for a rule of inference indicating how one formula may be permissibly transformed into another within the system. A common example of a rule used in this context is *modus ponens*:

$$\frac{(\phi \rightarrow \psi) \quad \phi}{\psi}$$

This rule tells us that from any pair of formulae that have the schematic forms represented by “ $(\phi \rightarrow \psi)$ ” and “ ϕ ” we can infer a formula with the form represented by “ ψ .” Again, there will be infinitely many pairs of formulae that have this form, and so the modus ponens schema allows us to generalize over all of them by finite means.

In more powerful logical languages and systems, such as the first-order predicate calculus and axiomatic systems of first-order logic, axiom schemas are again used to specify an infinite collection of axioms by finite means. However, unlike with those of propositional logic, the templates and rules for axiom schemas in first-order logic also deal with expressions of types other than just the wffs in the language, in line with the greater expressive power of the predicate calculus. The following axiom schemas for the universal quantifier “ $\forall$ ” in first-order logic illustrate this difference:

$$\text{PC1: } (\forall \alpha (\phi) \rightarrow \phi(\beta / \alpha))$$

$$\text{PC2: } (\forall \alpha (\phi \rightarrow \psi) \rightarrow (\forall \alpha (\phi) \rightarrow \forall \alpha (\psi)))$$

$$\text{PC3: } (\phi \rightarrow \forall \alpha (\phi)) \text{ when “}\alpha\text{” is not free in } \phi$$

In PC1, the expression “ $\phi(\beta/\alpha)$ ” indicates a metalinguistic variable for a wff that is exactly the same as that instantiating the metalinguistic variable “ ϕ ” *except* that it may differ (at most) in having the term that instantiates the metalinguistic variable “ β ” substituted for all occurrences of the bound term instantiating the metalinguistic variable “ α .” PC2 is more straightforward, indicating that the universal quantifier binding a term distributes over the material conditional. Finally, PC3

tells us that all instances of the conditional “ $(\phi \rightarrow \forall \alpha (\phi))$ ” are axioms whenever the bound term instantiating the metalinguistic variable “ α ” does not occur unbound (*free*) in the wff instantiating the metalinguistic variable “ ϕ .” As such, we can see that axiom schemas for axiomatic systems of first-order logic require variables for expressions other than wffs, that is, they require variables for *terms* (where the terms include variables, constants, and function applications in the language of the predicate calculus). Relative to a first-order language with variable terms “ $x_0, x_1, x_3 \dots$,” constant terms “ $c_0, c_2, c_3 \dots$,” (unary) function terms “ $f_0, f_1, f_3 \dots$,” and (monadic) predicate terms “ $F_0, F_1, F_3 \dots$.” Some possible instances of PC1–PC3, respectively, are the following:

$$\begin{aligned} &(\forall x_0 (F_1 x_0 \rightarrow F_1 x_0) \rightarrow (F_1 c_5 \rightarrow F_1 c_5)) \\ &(\forall x_7 (F_2 x_7 \rightarrow F_8 x_7) \rightarrow (\forall x_7 (F_2 x_7) \rightarrow \\ &\quad \forall x_7 (F_8 x_7))) \\ &(F_4 f_{88} (c_{17}) \rightarrow \forall x_3 (F_4 f_{88} (c_{17}))) \end{aligned}$$

As with the language of propositional logic, there will be infinitely many such axioms, most of which will be much more complicated than these three. The value of these axiom schemas is that we don’t have to worry about the impossible task of specifying these axioms individually, since the schemas allow us to capture them all in one fell swoop.

Axiom Schemas in Mathematics

Moving from logic to mathematics, axiom schemas are also indispensable in the formal study of axiomatic theories. However, it is perhaps worth noting that the techniques and practices of most working mathematicians are typically less emphatically formal than those of logicians. Two well-known examples are worth considering, however. The first is from first-order Zermelo–Fraenkel set theory and is known as the *axiom schema of separation*. One possible formulation of this axiom schema is as follows:

$$\forall s_0 (\exists s_1 (\forall s_2 ((s_2 \in s_1) \leftrightarrow ((s_2 \in s_0) \wedge \phi))))$$

Here, the metalinguistic variable “ ϕ ” can be replaced by any formula in the language of first-order set theory. In practice, “ ϕ ” will typically be instantiated with a formula in which the variable “ s_2 ” occurs free. Hence, the formula instantiating “ ϕ ” will represent some condition that the values of “ s_2 ” must satisfy, thus *separating* out from the sets that are the values for the variable “ s_0 ,” those subsets whose members all satisfy the condition in question. For example, an instance of the axiom schema of separation is

$$\forall s_0 \exists s_1 \forall s_2 \left((s_2 \in s_1) \leftrightarrow \left((s_2 \in s_0) \wedge \forall s_3 \left((s_3 \in s_2) \rightarrow \sim \exists s_4 (s_4 \in s_3) \right) \right) \right)$$

This instance of the schema is an axiom telling us that there is a (possibly empty) subset of every set where the members of that subset are such that their own members do not have any members themselves.

Another indispensable axiom schema shows up in formal theories of arithmetic. This is the axiom schema of induction, telling us that if zero satisfies some condition—that is, the value of the metalinguistic variable “ ϕ ”—and any number satisfying this condition implies that its successor *also* satisfies the condition, then all numbers satisfy the condition

$$\left(\phi(0) \rightarrow \left(\left(\forall x_0 (\phi(x_0) \rightarrow \phi(x_0 + 1)) \right) \rightarrow \forall x_0 (\phi(x_0)) \right) \right)$$

For example, here is one instance of the axiom of induction where the condition specified by the wff instantiating the metalinguistic variable, “ $\phi()$,” is the condition $\sum_{i=0}^{()} (i) = \frac{i(i+1)}{2}$ (I have omitted the numerical subscripts on the bound variable “ x ” for readability):

$$\left(\sum_{i=0}^{(0)} (i) = \frac{i(i+1)}{2} \rightarrow \left(\forall x \left(\sum_{i=0}^{(x)} (i) = \frac{i(i+1)}{2} \rightarrow \sum_{i=0}^{(x+1)} (i) = \frac{i(i+1)}{2} \right) \rightarrow \forall x \left(\sum_{i=0}^{(x)} (i) = \frac{i(i+1)}{2} \right) \right) \right)$$

Again, the impossible task of individually specifying every numerical property that we might want to deploy in an axiom of induction is sidestepped thanks to the use of an axiom schema.

Concluding Remarks

Throughout this entry, the notion of a metalinguistic variable has been ubiquitous. Indeed, the distinction between the *object language* (e.g., the formal languages of propositional logic, the predicate calculus, and first-order set theory) and the *metalanguage* (here, the language of English supplemented with symbols like the Greek letters used as variables) is arguably the essential technical distinction underlying the use of axiom schemas. The point of this distinction was made vividly by Tarski (1983, chap. 8) in his proposal of the *T-schema*, an axiom schema for axiomatizing the concept of *true sentence* in a formal language:

For any ϕ , “ ϕ ” is true if and only if S

Here the corner quotes surrounding “ ϕ ” indicate that we are *not* talking about what it takes for “ ϕ ” to be true, as this symbol is only a metalinguistic variable over sentences in the object language, not an actual sentence in the object language. Meanwhile, S is a variable representing the *translation* of the value of “ ϕ ” into the metalanguage. As Tarski helped to make clear, the distinction between the object language and metalanguage is needed to avoid *semantic closure*, in which the language has expressions that refer to other expressions within the same language. As MacFarlane makes clear on pages 61–65 in his textbook, avoiding this kind of closure, and thus avoiding the so-called *semantic paradoxes*, is another primary motivation for making use of axiom schemas in logic and mathematics.

Richard J. Teague

See also Formal Sciences; Logic and Language; Mathematics, 20th Century; Theories, Semantic Conception of; Set Theory

Further Readings

Burgess, J. P. (2022). *Set theory*. Cambridge University Press.

- Button, T., & Walsh, S. (2018) *Philosophy and model theory*. Oxford University Press.
- Corcoran, J. (2006) Schemata: The concept of schema in the history of logic. *Bulletin of Symbolic Logic*, 12(2): 219–240.
- MacFarlane, J. (2021). *Philosophical logic: A contemporary introduction*. Routledge.
- Tarski, A. (1983). *Logic, semantics, metamathematics: Papers from 1923 to 1938*. (Edited with introduction and analytic index by J. Corcoran.) Hackett Publishing Company.

AXIOMATIC THEORY

In ancient Greek philosophy, axioms were first principles that required no reasoning or justification as they could be believed to be immediately apparent. In modern philosophy of science, at least in the *non-statement view* of metascientific structuralism, axioms are part of a formal reconstruction of a theory which, according to the founders of this movement in philosophy of science, abated the longish discussion in the community of the translatability between observational and theoretical terms and statements. This entry traces the development of axiomatic theory from its origins, describes the elements of the theory and its models, and concludes with a note on its applications in various sciences.

Observational and Theoretical Statements

Rudolf Carnap, in his *Philosophical Foundations of Physics*, still found “the distinction between what may be called . . . empirical laws and theoretical laws” [. . .] “one of the most important distinctions between two types of laws in science,” empirical laws referring to “properties directly perceived by the senses” such as “blue,” “hard,” “hot” from the point of view of the philosopher, or properties “that can be measured in a relatively simple way” (Carnap, 1966, p. 225). But the distinction between properties perceivable by the senses and those measured in a *relatively* simple way is not really clear-cut, and the simplicity of a measurement is not clearly definable either. Moreover, the problem persists that

theoretical laws were formulated in theoretical terms, whereas empirical laws were formulated in observable terms.

Hence, it became necessary to introduce so-called correspondence rules for a translation between the two kinds of terms, for instance between the temperature of a gas as an observable in Carnap’s wider sense—temperature being a term of the observational language—and the mean kinetic energy of its molecules as a theoretical term in the theoretical language, or to give an example from the social sciences, between the answer to a question such as “How much confidence do you have in President Joseph Biden?” with possible answers “a lot of confidence, some confidence, not too much confidence, no confidence at all” as an observable and the propensity to cast a vote in favor of President Biden, the latter being a theoretical term, because not even the interviewee would be able to give a numerical value to their probability to vote in the oncoming election for the incumbent candidate. The latter example shows that what is often called the operationalization of a theoretical term can be merely arbitrary with the help of one or more such observable answers to a related question. Another collection of possible answers, for example, 11 or 10 or 5 instead of 4 possible answers, would lead to another operationalization result. To keep this example, one could even argue that one needs some other theory relating the propensity to vote for a certain candidate to the propensity to give certain answers to certain questions, such that correspondence rules would be necessary to link attitudes, voting, and answer propensities to answers given in a certain interview situation.

The so-called syntactic view of theory structure uses these correspondence rules to *translate* between observational and theoretical language (but this kind of *translation* is obviously not of the same kind as a translation between, say, English and Chinese).

The first but incomplete attempt to overcome the dichotomy between observable and theoretical terms was made by Frank P. Ramsey in what is called the *Ramsey sentence of a theory* (also referred to as a *Carnap sentence*), which eliminates the theoretical terms from a theoretical law. If the interpreted theory

$TC = T(\tau_1, \dots, \tau_n) \wedge C(\tau_{i_1}, \dots, \tau_{i_k}; \omega_1, \dots, \omega_m)$ is the conjunction of all theoretical postulates T (all of them formulated in the theoretical language) and all necessary correspondence rules C (these in turn using terms from both languages, observational and theoretical), then it is possible to replace the theoretical terms τ_{i_j} which occur in the correspondence rules (with its observables $\omega_1 \dots \omega_n$) with new variables ϕ_i and replace the interpreted theory with its Ramsey sentence, which postulates that $\exists \phi_1 \dots \exists \phi_k TC(\phi_1, \dots, \phi_k, \omega_1, \dots, \omega_n)$ now claims that substitutions exist for the uninterpreted theoretical terms.

This elimination, however, leaves open the question of how theories are linked together, as in the voting propensity example above. The solution came with the so-called semantic view of theory structure, which introduced set-theoretical entities instead of linguistic ones; this was John D. Sneed's effort to define the logical structure of a theory with set-theoretic predicates instead of speaking about statements whether they are observational or theoretical (which is why this view on theories is called the *non-statement view*).

Elements of the Theory and Its Models

The structuralist program, which builds upon Sneed's work, is centered on the definition of theory elements. A *theory element* is a mathematical structure consisting of a theory core and a set of intended applications, whereby the theory core describes what can be said about the intended applications of the theory core. Theory elements are usually one or a few laws within a scientific discipline, for instance Hooke's spring law (HSL) or classical particle mechanics dealing with masses, forces, velocities, and so on, as the origin of structuralist consideration was clearly within physics, mainly because in physics mathematical formalizations have a long tradition dating at least as far back as Newton and Kepler.

The theory core is a set-theoretic predicate which is typically defined as follows:

T is a theory element if and only if there exist sets of partial potential models M_{pp} , of potential

models M_p , of models M , of constraints GC , of links GL and of intended applications I such that

$K = \langle M_{pp}, M_p, M, GC, GL \rangle$ is a theory core,

$T = \langle K, I \rangle$ is the theory element and

$I \subseteq M_{pp}$

The set of partial potential models M_{pp} is defining its elements x as another mathematical structure listing all terms which are necessary to speak about the theory in question. In the case of Hooke's law, this list contains a finite set of springs S , a finite set of identical pieces of metal or some other material which can easily brought into an identical form (*weights*) W , three functions $n(w, s)$, $l(w, s)$, and $k(w, s) = l(w, s)/n(w, s)$ yielding the number of weights in a collection w of these weights hanging from a spring and the length of the extension of the spring caused by the weight collection w , whereas $k(w, s)$ is the outcome of an experiment with spring s and weight collection w . In this simple example, the difference between M_p and M_{pp} is very small; it consists of the observation that for small lengths of the extension of the spring, k does not depend on w . The terms s , w , n , l and k can each separately be defined without knowing anything about mechanics: Springs and weights can be identified as such, weights in a collection can be counted, extension lengths can be measured, and the quotient k can be calculated. It is only the assumption that k does not depend on the size of the collection of weights but only on the characteristics of the spring that allows for a redefinition of $k(s) = l(w, s)/n(w, s)$ and separates partial potential models from potential models, as the full model is a potential model for which the axiom of Hooke's spring law holds, namely that for all collections of weights and for each spring which is not extended beyond a certain maximum extension $l_0(s)$, $k(s) = l(w, s)/n(w, s)$ is constant and independent on the collection of weights, from which it follows that $l_0(s)$ and $k(s)$ can be considered as device constants. Moreover, as the experimenter will perhaps use their hands to extend the spring, they will feel Carnap's property directly perceived by the senses as it costs more and more effort the longer the extension is. Hence, the

experimenter will assume that there is a monotonically increasing function between their effort and the extension, and it seems reasonably to define this function as a linear function—the HSL-theoretical term *force*—and use it to make this effort quantitatively measurable as proportional to the length of the extension.

The difference between observable terms and theoretical terms is redefined in a straightforward manner as *theoreticity*, always in reference to a theory in question and not an absolute feature of a term. Counting weights, measuring lengths, and dividing numbers necessitate their own theories but not a theory of springs and weights. Hence, the force and the device constants $l_0(s)$ and $k(s)$ are HSL-theoretical terms, all the others are not (and perhaps the value of $l_0(s)$ depends on the accuracy with which $l(w, s)$ and consequently $k(s)$ is measured; the precision of $n(w, s)$ is a lesser problem in this idealized experiment).

Constraints, Links, and Theory Nets

Constraints of a theory collected in the set *GC* are about restrictions to its applicability; in the simple example above, this could be the constraint that all pieces are of the same volume, form, and of the same chemical composition (i.e., the theory would not apply to pieces some of which are balls whereas others might be extremely flat but with the same volume) and that these three features are constant over time (i.e., the theory would not apply to identical pieces of ice which might melt away over time).

Links of a theory element collected in the set *GL* connect it to other theory elements; in the simple example above, volume and chemical composition are used as if they were properties “that can be measured in a relatively simple way.” But to determine the chemical composition of a weight hanging from a spring, some chemistry is necessary; and even to determine the equality of the volume of several metal balls it is necessary to measure their diameters in cm and to assume that the same formula from geometry yielding the volume in cm^3 applies to all of them. Hence, in a way the theoretical links replace the correspondence rules of earlier approaches, but with the important difference that the observational part of a correspondence rule is free from

any theory. Links are defined as sets of tuples of terms which connect correspondent terms of the two theories in question. In the example above, there is the term l denoting the length of the extension of the spring in the theory of Hooke springs and the term l^* in geometry, both of which map to integer numbers (multiples of a certain small unit length).

Not only can two theories be linked together, but theory nets can be built that potentially encompass wide areas within and across disciplines. Thus, the discussion about reducing an *observational law*—say the law $p\nu = kT$ connecting pressure p , temperature T , and volume ν of a gas—can be restated in terms of a theoretical law

assuming that $p = \frac{2E}{3\nu} = kT$, where the pressure is

related to the mean kinetic energy E of the gas molecules. This also ends the discussion about *reductionism* as the assumption that social science could somehow be *reduced* to psychology, the latter to neurobiology, and further to chemistry, and still further to particle physics; instead theoretical terms of a theory in one of these disciplines could be linked to (instead of reduced to) theoretical terms of another discipline.

In the social sciences, most theories have many more terms than in the simple example from classical mechanics. Nevertheless, a structuralist reconstruction of middle-range theories is possible, too. A theory, call it AOF, about attitude or opinion formation, containing terms such as the propensity to vote for a certain candidate or party and the influence of communicating attitudes between citizens certainly has a number of AOF-theoretical terms—for example, $\alpha: P \times T \rightarrow R$ such that $\alpha(p, t)$ is the position of person p at time t on a left–right scale and $\delta: P \times P \rightarrow R$ —such that $\delta(p_1, p_2)$ and $\delta(p_2, p_1)$ are the movements of Person p_1 and Person p_2 on the left–right scale after they met each other and they have discussed their attitudes toward the candidates and parties running in an oncoming election. In an application of such a theory, the distributions of votes cast for candidates and parties would be observable as AOF-non-theoretical terms (as these could easily be counted) and the distribution of answers to questions like the one cited above (“How much confidence do you have in President

Biden?”). A theory about measuring attitudes with algebraic algorithms such as principal component analysis (PCA) could be linked to AOF, and this intertheoretic link would help to reinterpret the PCA-theoretical terms α and δ as AOF-non-theoretic terms. The function that connects $\delta(p_1, p_2)$ on one hand with $\alpha(p_1, t)$ and $\alpha(p_2, t)$ on the other hand would then be the actual AOF-theoretical term. The theory would have to provide axioms that govern the development of the individual attitude changes and of the distribution of attitudes over a large population in which persons meet each other under certain conditions. Some of these conditions of meeting might be *observable*, in other words AOF-non-theoretical, whereas other such conditions—for instance preferences to discuss with other persons depending on the known or guessed attitude difference between two persons who meet occasionally cannot be directly observed. A few attempts to reconstruct theories of opinion formation have already been made and show the difficulties of such attempts, which are mainly caused by at least two of the big differences between the systems physical theories are about and those which social scientists deal with: The heterogeneity of the *particles* social scientists have to take into account and the fact that human beings, unlike physical particles, can communicate individually and change the internal state of other individuals. Hence, theory nets in economics, and the social sciences generally, but also in the environmental sciences with their extremely complex systems will be much more widespread than in physics. (Balzer, Moulines, and Sneed, the founding fathers of the structuralist program, in their 1987 book themselves reconstructed only pure exchange economics with an intended application that they described as “some specific village at some specific time” (Balzer et al., 1987, p. 161) and did not continue their approach to the formulation of constraints or links.)

Applications in Various Sciences

Since the late 1970s, the method of formally reconstructing theories according to the *non-statement* view has been applied to various (pre-formal) theories in various disciplines. As the movement started within physics, the first candidates to be

reconstructed were taken from classical physics, thermodynamics, quantum mechanics, chemistry, biology, medicine, neuroscience, psychology, economics, management, and the social and political sciences. It is not surprising that the density of publications in these disciplines decreases along this list, and it is particularly striking that the acceptance of the *non-statement view* seems to have been greater in German- and Spanish-speaking countries than in English-speaking countries (although English as a publication language still prevails). The disciplines at the end of the list above are much less formalized and quantified than those at the beginning of the list; hence the late and sparse introduction of theory reconstruction into studies of organizations and large social and political systems is understandable, as the reconstruction of a theory that was formulated mathematically long ago is much easier than that of a verbally formulated theory with all its ambiguities. With the emergence and continued development of computational social science, this situation will certainly change in the years to come.

Klaus G. Troitzsch

See also Data Models; Logical Theory, Structuralism

Further Readings

- Abreu, C., Lorenzano, P., & Moulines, U. (2013). Bibliography of structuralism III. *Metatheoria—Revista de Filosofía e Historia de la Ciencia*, 3(2), 1–36.
- Abreu, C., & Lorenzano, P. (2020). Conversaciones con C. Ulises Moulines. *Metatheoria—Revista de Filosofía e Historia de la Ciencia*, 11(1), 1–31.
- Balzer, W., Moulines, C. U., & Sneed, J. D. (1987). *An architectonic for science. The structuralist program*. Reidel.
- Carnap, R. (1966). *Philosophical foundations of physics: An introduction to the philosophy of science*. Basic Books.
- Nagel, E. (1961). *The structure of science. Problems in the logic of scientific explanation*. Routledge and Kegan Paul.
- Sneed, J. D. (1979). *The logical structure of mathematical physics*. Reidel.
- Stegmüller, W. (1979). *The structuralist view of theories. A possible analogue of the Bourbaki programme in physical science*. Springer.

B

BELIEF REVISION

The stability and malleability of belief is the concern of belief revision. As with most theoretical language, belief and its revision have multiple meanings depending on purpose, context, and logic, so any case must define its terms and their scope. This entry focuses on belief revision within the context of scientific reasoning.

Stakeholders are divided by whether they orient belief in the realm of cognitive (knowledge and reasoning) or affective (emotions, attitudes, values, and spiritual belief). The topic of belief revision need not address the stability or malleability of the world one is believing in but simply the dynamics that influence belief persistence, adjustment, or abandonment. Although individuals may revise beliefs pragmatically, beliefs in the sense of scientific theory or a computer software's logic would require a deliberate strategy reflecting a philosophy of science. For instance, Karl Popper (1902–1994) would reject a prototype model if any part diverged, but Willard Van Orman Quine (1908–2000) would revise or *fine-tune* the prototype until all parts became congruent.

Defining Belief and Its Revision

The revision function depends upon the defined scope of the belief, the logic used to determine doubt or confidence, and the epistemic aims and values regarding the belief. Belief may represent *propositions of truth*, justified by rational argument

to produce knowledge. Belief may also suggest the degree of confidence in the proposition (faith) or it may provide the justification for other beliefs (theory). Belief may denote a pattern of interpretation (habits of mind) or judgment (ethics and morals). A belief system is one in which truths form a coherent framework of understanding (ideology); that is, an integrated, coherent protocol for making decisions (priorities), predictions (hypotheses), or solving problems (algorithms).

Theories Explaining Belief Revision

The first definition above is the standard used in *scientific reasoning* and especially in computer programming. Each coded statement contributes a belief to the mechanical logic. These statements reduce the complexity of a scenario to very small increments, forming a cohesive set of responses to new information. Advanced programs are called *artificial intelligence* if they *learn*, that is, if they can change their own coding. Belief revision thus refers to advanced coding logic for designing computational programs. Apart from the technicalities involved, this discourse informs analytic philosophy, which will symbolize propositions in transformational language akin to algebra.

If each belief serves to scaffold further beliefs, there is a foundation required to reach complex heights. The *belief set* is a closed set of propositions under examination, while *belief base* can examine separate propositions that are elements of the base and can tell the difference between those that are explicitly stated and those that are

a logical extension of the explicit beliefs. The primary goal of such a *foundation theory* is to discover truth, and thus a faulty premise casts doubt on the entire code sequence. This coding logic will identify an erroneous proposition and then either accept it (expansion), remove it (contraction), or both add something and take something else away (revision). It is a syntactic approach, akin to using syllogism for deductive reasoning. By contrast, a *coherence theory* is a synthetic approach, accepting new propositions for their logic in maintaining the meaningfulness of the whole framework; in other words, *verisimilitude*. Stability of the network of decisions is more important than semantic integrity, so the existing belief structure is not revised so much as its applications. A computer program example would be the apparently intuitive helpfulness of online search engines to provide advertisements consistent with topics related to your own history of links. The purpose of the algorithm is to generate all links consistent with what appears to be your consumer interest; no ideology is at risk for the computer by adding more search items. As long as the additional information is consistent with your own belief about the purpose of being online, you do not revise your belief; however, once the cost of these additions is too great, such as when the ads are distractingly animated and intrusive, you may revise your belief. According to foundation theory, you might abandon that browser or online shopping if you conclude the practice is wrong, but according to coherence theory, you might simply work around the ads if they are tolerable, or the stability of your habit takes on greater value than the dissonance of incompatible behavior.

Experimental studies of belief revision may be structured in series with multivariate analyses used to identify effect sizes for a range of possible influences. The formulas used in the statistical packages are themselves the source of inquiry now that computation can occur so quickly. Repeated measures of cognitive learning outcomes in schools are used to monitor cognitive change interpreted according to pedagogical theory.

Belief Persistence

One reason a belief is held in the face of contrary evidence is simply habit for individuals and the

momentum of a complex network for organizations. Habits are automatic, whereas belief revision requires deliberation. There is also an emotional dimension of enjoying the predictability of customary beliefs and not trusting that the work and discomfort of extinguishing the old ones will be adequately compensated by a better future. Yet another reason beliefs do not change is that memories are faulty at best, it being impossible to retrieve all details of arguments that lead to confidence in the incumbent decision. Also, individual belief systems are fraught with inconsistencies, and people often will tolerate beliefs that do not seem to be aligned. For instance, one can wholeheartedly agree with the need to buy from independent local businesses but will nonetheless search out a bargain at a distant warehouse store.

Many beliefs about the world are formed intuitively at a young age. Teachers must diagnose their students' knowledge base, including misconceptions, in order to design instruction that prompts a more enlightened understanding and assessment that measures the degree of confidence in the new perspective. In this way, learning can be construed as a revision of belief. A particular challenge is posed when the academic objective is dispositional; that is, when students are expected to become more tolerant, frugal, democratic, generous, or healthy. Meanwhile, the aforementioned factors of habit, memory, comfort, and purpose conspire to maintain false beliefs or disbelief.

Belief profiles have been found to correlate with patterns of behavior and conceptual change, and thus belief revision is of interest to those hoping to influence change as well as those simply curious to understand the phenomenon. Given their individual conventions regarding the concept of belief, there are different fields of inquiry concerned with belief revision, for example, education (nature of teaching and learning), psychology (existential crises; self-efficacy), religion (conversions; hermeneutics), business (market analysis; organization management), mathematics (proofs; inferential statistics), science (empirical justification), and computer programming (artificial intelligence).

Many theories have emerged out of interest in belief revision. Thomas Kuhn (1922–1996) described a *paradigm shift* after which a revolutionary idea is

widely assumed to be true, for example, evolution or climate change, each of which has substantial empirical evidence to support it. The correlation for personal beliefs would be the loss of faith after some traumatic event, for instance, one's belief in God or one's trust in banks or one's commitment to marriage. The result is an existential crisis, a conversion to a new belief system, or perhaps agnosticism; that is, not knowing what is true, or perhaps an antagonistic stance such as atheism. Holocaust survivor Viktor Frankl (1905–1997) developed *logotherapy* to explain the sudden deaths of prisoners even when circumstances had not changed. He concluded that all people search for meaning in their lives, a useful construct in mental health.

Cybernetics is a system theory concerned with the dynamics of interdependence. While associated with mechanical or nonhuman systems, the principles of interaction apply to social contexts, especially regarding feedback, whether cognitive or affective. Conflict resolution techniques are a result, making the case for interrupting cycles of violence in any scale, such as within families or between countries. In such matters of diplomacy and legality, terms of formal agreements between parties will specify behaviors associated with target beliefs about their identities and status, given that in any social system there is a continuously challenging landscape of problems with sharing resources and power. Both what is believed and how it is believed will influence the stability of the relationships which are maintained on the basis of continual interaction with trusted sources of information. Open communication is blurred by tactics of silencing voices and withholding or distorting information and of limiting sources to an *echo chamber* that accepts only that which conforms to one's belief.

Naomi Jeffery Petersen

See also Fact Versus Theory; Generalization; Mental Models; Theory Change

Further Readings

- Godden, D. M. (2012). Rethinking the debriefing paradigm: The rationality of belief perseverance. *Logos & Episteme: An International Journal of Epistemology*, 3(1), 51–74
- Hansson, S. O. (2014). Logic of belief revision. In E. N. Zalta (Ed.), *The Stanford encyclopedia of philosophy* (Spring 2014 ed.). <http://plato.stanford.edu/archives/spr2014/entries/logic-belief-revision>
- Lane, J. D., & Paul, L. H. (2014). Confronting, representing, and believing counterintuitive concepts: Navigating the natural and the supernatural. *Perspectives on Psychological Science*, 9(2), 144–160.
- Lee, B. D. (1998). The paradox of belief instability and a revision theory of belief. *Pacific Philosophical Quarterly*, 79(4), 314–328.
- Mason, L. (2010). Beliefs about knowledge and revision of knowledge: On the importance of epistemic beliefs for intentional conceptual change in elementary and middle school students. In L. D. Bendixen & F. C. Feucht (Eds.), *Personal epistemology in the classroom: Theory, research, and implications for practice* (pp. 258–291). Cambridge University Press.
- Schlottmann, A., & Norman, H. A. (1995). Belief revision in children: Serial judgment in social cognition and decision-making domains. *Journal of Experimental Psychology: Learning, Memory & Cognition*, 21, 1349–1364.
- Urban, T. (2019). The thinking ladder. *Wait But Why*. <https://waitbutwhy.com/2019/09/thinking-ladder.html>

BIG DATA

Our current age is often referred to as the era of *big data*. While the term has become increasingly popular in the last 10 years, the concept is now used in a wide variety of ways, often without referring to an acknowledged technical definition of the concept. In fact, an agreed-upon and technical definition can be argued not to exist, and the term *big data* is mainly used to refer to broad and general tendencies related to a rapid increase in the data collected in various domains and the accompanying need to develop and employ new methods of analysis in order to make sense of such data. These methods of analysis connect big data intimately to data mining, artificial intelligence (AI), and machine learning.

Big data has become a ubiquitous phenomenon, and it is applied to domains such as research, business intelligence, social networks, health care, crime prevention, and anti-terrorism. Data is routinely described as the “new oil,” and in parallel with the historical discovery of other precious

resources, the quest to drill, mine, and refine it is both intense and characterized by competition, rapid development, and private sector innovation. Big data carries a number of actual and potential benefits, but it is also related to a number of ethical challenges. Many are related to the nature of the data employed, some relate to a purported blind and dangerous faith in quantitative data and algorithms as the solution to and source of insight in *all* aspects of human lives, and some are mainly related to the use of algorithmic decision-making in general. The next section of this entry presents an expanded definition of big data, followed by an exploration of the philosophical basis of the concept and its roots in the quantitative and empiricist approach to knowledge. Before concluding, the entry offers a discussion of the ethical considerations surrounding the growing applications of big data.

Definition

Big data refers to a development toward generating and gathering increasing amounts of data at increasing speeds, while the data is structured in heterogeneous data sets. The data in question can, for example, be social and personal, or related to industrial processes or scientific measurements, and several sources and types of data are often connected in large data sets.

Although *data* is of course a key component of big data, the tools and techniques used to store, manage, and analyze the data are also integral parts of the concept. The problem with the data sets handled in big data is that traditional methods of storage and analysis are insufficient. Distributed storage systems and cloud computing are used for storing and managing the data. Data mining and machine-learning-based AI have become terms intimately associated with the analysis of the large data sets. The entire chain of generating, gathering, storing, and analyzing data is included as part of what is referred to as *big data*.

Data in isolation—gathered and stored—has limited potential for positive or negative consequences. It is through the quest to derive value from the data that, for example, benefits in business intelligence and concerns over discriminatory practices arise. The term *big data* is now used by

both data and computer scientists, social scientists, and the general public, and the latter two groups often refer to a general and nonspecified concept without referring to particular forms of database management or specific machine learning techniques.

The techniques used to analyze the data are not entirely new, and these stem from earlier academic work in AI and statistics. However, the volume of today's data sets has mainly become a reality through private corporations focused on gathering and using data to provide various services, and this has led to a situation in which much development is done outside of academia. Although private companies are important for developing big data techniques, government and other public institutions are important secondary beneficiaries. There is now a growing tendency for government agencies to use big data in order to, among many other things, prevent crime and terrorism, find effective pandemic responses, and both discover fraud and automate decision-making in social services.

Philosophical Foundation

As data is analyzed in order to answer certain questions, or achieve a set of goals, some philosophical assumptions are inevitably involved. These *can* be made explicit, but far more often, as big data is applied to solve issues of importance for businesses and society, the assumptions remain implicit. One reason is that a lot of the techniques are developed and applied in nonacademic settings, but another important reason is that proponents of big data have repeatedly argued that big data is *not* based on theory or ideology. One of the major causes of disagreement over the implications of big data revolves around the possibility of achieving neutral and objective science through analyses of raw data and partially autonomous AI.

Scientific and other applications of big data are characterized by a multitude of approaches, and in the following a *typical* (but not the only possible) approach is described. By removing the human scientist from the equation, it is argued, true objective science has finally been realized. Such claims disregard the nature and origin of the data, and that data in itself is not neutral. The

various choices related to how data is gathered, which variables are collected, how data is coded, and so forth, by necessity involve a set of choices resulting in data that is a *model* of reality and not an accurate reproduction of it. Also, human beings are heavily involved in the creation of the algorithms used in the analysis of data, decisions about when and how to use big data solutions, and also in choosing how the results are interpreted and utilized.

Big data is based on a quantitative and empiricist approach to knowledge, in which observable phenomena are seen as the basis of knowledge. As machine learning methods have become increasingly advanced, some even argue that human interpretation is no longer an integral part of deriving insight from these observations. Furthermore, *more* data is seen as conducive to more knowledge, whereby observations both strengthen the results of the analysis *and* improve the tools of analysis as data is used to train and further develop machine learning algorithms.

As data is gathered and fed to the computer for analysis, the intent is that AI tools sift through it and uncover interesting patterns in the data which are not identifiable directly by humans or through traditional statistical methods. This process is referred to as *data mining*. Big data is thus usually characterized by inductivism and the idea that we blindly approach data in order to identify patterns and develop theories about dependencies and causation. In practice, however, human beings are necessarily involved in a dynamic process in which initial results from the analysis are examined and used to guide and shape further analysis. Such a process would be more accurately described as a combination of the deductive and inductive approaches, at times referred to as an *abductive* or *retroductive* approach.

Big data as applied to human phenomena is often based on a behaviorist approach, in which observable phenomena and actions are emphasized over, for example, cognitive phenomena and interpretation. When an individual's subjective experiences and evaluations constitute the data analyzed, the goal is likely to be how this translates into actions, and how various variables are connected, rather than a deeper understanding and explanation of what goes on inside people's heads.

In short, big data is philosophically closer to the natural sciences than the social sciences and humanities. The knowledge and insight produced by big data can thus be argued to constitute a part of the answer when human phenomena are examined, and not a complete picture. In order to explain and understand *why* the patterns discovered occur, and how the phenomena and correlations described are actually experienced, human scientists, other methods of analysis, and other forms of data are still vital parts of making sense of the world as it appears through the results produced through big data.

Ethical Considerations

Big data is a major part of much contemporary debate about surveillance, privacy, social networks, and liberty. Although not directly related to big data as *large data sets*, such debate is clearly related to some of the more prominent applications of the general phenomenon of big data understood as large data sets combined with AI for analysis. In this section, the discussion narrows in on *social* and *personal* big data, since data related to individuals and their social relations is the most obvious cause of concern.

A key concern related to big data is a purported lack of transparency and difficulties of explaining the results of AI-based analysis of data sets. Many of the purported benefits of big data are incontrovertible, but an opaque process of analysis often likened to a "black box" creates a situation in which unintended consequences often go unnoticed. For this reason, much research on AI is now focused on creating *explainable AI* because this is required for securing transparency and ensuring that big data and AI are developed and deployed in an accountable and responsible manner.

Because big data is often based on historical data, its use can perpetuate and even obfuscate structural racism, sexism, and other forms of inequality. For example, if a bank employs an algorithm for determining who should receive a loan based on historical data, the models trained on this data will perpetuate historical bias related to, for example, race and gender. Furthermore, although we know that human beings are biased, some argue that machines are *not* biased; therefore, unless the algorithms are audited and

developed in order to foster explanation, discriminatory practices may go unnoticed and even become legitimized.

Privacy and surveillance are other concepts often related to big data, as the generation and gathering of data can be perceived as a form of surveillance. Concerns relate to a violation of rights to privacy in general, and to more insidious phenomena, such as the combination of large numbers of data points from different databases in order to create highly detailed personality profiles. This allows actors to influence individuals more effectively than they could otherwise have done and could thus be construed as inimical to liberty and independence. Such influence relates to all sorts of behavior, including political behavior and government efforts to exert social control.

Whereas the aforementioned use of personal data often involves efforts to market and sell products and services more effectively, the societal consequences of big data are also a cause for concern. Particularly in relation to social media, some authors have shown how big-data-based personalization might allow actors to influence individuals' perceptions of the world (trap them in "filter bubbles" where people are increasingly being exposed to what they *want* to see), influence how people associate and relate to each other (and create "echo chambers" in which they associate only with like-minded people), and thus contribute to societal polarization, which is a growing concern.

In contrast to the preceding challenges, the benefits derivable from big data can potentially help improve individuals' lives and alleviate and solve important social problems. For example, preventing crime and terrorism is clearly a good thing and can easily be argued to warrant the sacrifice on some privacy. The same applies to using big data in health care, where it is used to improve the understanding of diseases and interventions and for more effective diagnosis. The purpose of highlighting the causes for concern mentioned above is simply to show that while big data is highly effective and provides a plethora of new opportunities in all domains of society, it is important to include perspectives from ethics and the social sciences in the development and deployment of big data projects.

Conclusion

While big data is arguably a continuation of a long quest for more information and better methods of analysis, recent developments related to the massive amounts of data gathered from various devices such as household appliances, smart electricity meters, global-positioning-system-enabled devices and other devices with a variety of sensors, and social media suggest that it is reasonable to describe Big Data as a new and important phenomenon. The amount of data has increased rapidly, and even if the machine learning techniques used are not entirely new, they are now trained on far larger data sets than before, which leads to increasingly useful results and a renewed interest in research and development.

Big data is used in all domains of modern societies, and the benefits it provides are substantial. These include, among other things, benefits in terms of scientific advancements, improved business intelligence and analytics, new tools and methods for law enforcement, innovation in health care, and new ways for people and societies to connect, communicate, and share information.

At the same time, there are important ethical concerns raised by big data, which include invasions of privacy, the non-neutrality of data and discrimination, societal effects of social networks, and the use of big data to exercise political influence and control. Those in possession of large data sets gain power and the means to influence others more effectively than ever before.

Henrik Skaug Sætra

See also Artificial Intelligence; Data and Phenomena; Data Models; Philosophy of Science; Prediction

Further Readings

- Boyd, D., & Crawford, K. (2012). Critical questions for big data: Provocations for a cultural, technological, and scholarly phenomenon. *Information, Communication & Society*, 15(5), 662–679. <https://doi.org/10.1080/1369118X.2012.678878>
- Chen, M., Mao, S., & Liu, Y. (2014). Big data: A survey. *Mobile Networks and Applications*, 19(2), 171–209. <https://doi.org/10.1007/s11036-013-0489-0>
- Dean, J. (2014). *Big data, data mining, and machine learning: Value creation for business leaders and practitioners*. John Wiley & Sons.

- Kitchin, R. (2014). Big data, new epistemologies and paradigm shifts. *Big Data & Society*, 1(1). <https://doi.org/10.1177/2053951714528481>
- Mayer-Schönberger, V., & Cukier, K. (2013). *Big data: A revolution that will transform how we live, work, and think*. Houghton Mifflin Harcourt.
- Sætra, H. S. (2018). Science as a vocation in the era of big data: The philosophy of science behind big data and humanity's continued part in science. *Integrative Psychological and Behavioral Science*, 52(4), 508–522. <https://doi.org/10.1007/s12124-018-9447-5>
- Wu, X., Zhu, X., Wu, G. Q., & Ding, W. (2013). Data mining with big data. *IEEE Transactions on Knowledge and Data Engineering*, 26(1), 97–107. <https://doi.org/10.1109/TKDE.2013.109>
- Zuboff, S. (2019). *The age of surveillance capitalism: The fight for the future at the new frontier of power*. Profile Books.

BIOCHEMISTRY, 19TH CENTURY

The field of biochemistry has a long-standing history of immense contributions to the progress of medicine. Notwithstanding, biochemistry has attracted less attention than the other biomedical sciences. Much of this can be attributed to the fact that as a separate entity it is rather young when compared with the likes of anatomy or physiology. Before 1750, biochemistry was not a full-fledged theory in its own right. It did gain a measure of respectability as its own discipline, however, in the later part of the 19th century and would come to fruition thereafter.

Although a strong association can be traced between biochemistry and organic chemistry, the history of biochemistry has received scant coverage within the history of chemistry generally. This is perhaps because biochemistry was once regarded by chemists as a lesser subdiscipline or as a part of medicine. In its broader sense, biochemistry can be regarded as the study of living matter and its chemical composition as well as of the biochemical processes essential for the maintenance of growth and vital life activities. This entry aims to examine the various developments within the sphere of biochemistry in the 19th century that were important to its subsequent emergence in the 20th century. Significant contributions by eminent scientists, such as the discovery of the activity of

yeasts in alcoholic fermentation, discovery of the first enzyme, and the classification and nomenclature of the enzymes, were all crucial leads for the future generations of biochemists. In this phase, biochemistry was mainly concerned with the study of chemical reactions taking place within living cells (both in microorganisms and in humans).

Roots of Biochemistry: An Overview

The roots of biochemistry can be found in 19th-century studies of organic chemistry, animal chemistry, and physiological chemistry. It was already found by then that the chemistry of living and nonliving materials was markedly different. In the 1830s, it was thought that the jellylike homogeneous substance, protoplasm, present within the organisms was responsible for carrying out most of the vital activities of biosynthesis, respiration, and intracellular food breakdown (metabolism). However, this general belief did not find support from such eminent chemists as Justus von Liebig (1803–1873) or Ernst Hoppe-Seyler (1825–1895). Although the hydrolytic enzymes such as pepsin, maltase, and amylase were known to the scientific world at that time, they were not considered to be acting within the cells. Eduard Buchner in 1897 perhaps performed the single most significant experiment that commenced the study of biochemistry. He prepared a cell-free yeast extract which he named zymase and observed its efficiency in fermenting glucose into ethanol and carbon dioxide. He later went on to receive the Nobel Prize (in 1907) for his discovery of cell-free fermentation and contribution to biochemical research. Zymase was regarded by Buchner as a single enzyme; however, this was soon disproved by others and was found to contain several other enzymes. Nonetheless, Buchner's work discredited the protoplasm theory and confirmed fermentation as a chemical process. Also, the presumed differences between the catalytic activities of the intracellular enzymes and the extracellular hydrolytic enzymes were found not to exist.

The French chemist Louis Pasteur (1822–1895) started studying the alcoholic fermentation of sugar by yeasts. He went on to conclude that the fermentation of sugar to alcohol is carried out by a vital force called *ferments*, present within the yeast cells, that can only function inside the living

cells. He further went on to clarify that fermentation was not associated with the putrefaction or death of the cells but rather was correlated with the organization and life of the yeast cells. In 1833, the first enzyme, called *diastase*, was discovered by the French chemist Anselme Payen, and in 1878 the German physiologist Wilhelm Kühne coined the term *enzyme* to describe this process. Nonliving substances such as pepsin was later referred to as *enzymes*, while the term *ferment* was used to describe the chemical processes of the living organisms.

Anselme Payen and Diastase

Payen, the son of an entrepreneur who had started a number of chemical production factories, rose to fame with his discoveries of the carbohydrate cellulose and the enzyme diastase. He had received knowledge of chemistry from his father and later from the likes of Michel-Eugène Chevreul and Nicolas Louis Vauquelin. Payen was promoted to become the head of his father's borax production plant in 1815. Here, he was able to use boric acid (available cheaply from Italy) to prepare borax. With the production costs going down significantly, Payen succeeded in besting his competitors in selling borax. Payen then shifted his attention in studying the sugar production from sugar beets and thereby established a method of sugar decolorization using charcoal. Subsequently, charcoal filters began to be used in gas masks to absorb dangerous organic gases. In 1833, Payen came up with yet another discovery: a chemical derived from malt extract that could catalyze the conversion of starch to sugar. The catalyst was named diastase, the first enzyme to be produced in concentrated form. The pattern of naming the enzyme diastase, as initiated by Payen, is still existent today with enzymes being named with the suffix *-ase*. He also concentrated his research on wood and extracted from it a substance resembling starch. This substance he named *cellulose*, which he could find in abundance in the plant cell walls he studied. Here also Payen became the pioneer in starting a nomenclature of the carbohydrates that would end with the suffix *-ose*. Although Payen was the first to isolate cellulose, it was the American botanist Wanda K. Farr (1895–1983) who ultimately discovered the mechanism of its

production by plants. Many other subsequent inventions used cellulose as a building block. It became the main component in a number of products, including cellophane, celluloid, rayon, collodion, guncotton, nitrocellulose, and other related products.

Louis Pasteur and Alcoholic Fermentation

The first of Pasteur's contribution to the study on microbial activities was in relation to lactic acid fermentation. Pasteur was consulted by a local alcohol producer, a Monsieur Bigo, who was facing serious problems with fermentation. After careful examination of the fermenting liquor under a microscope, Pasteur found that when the fermentation was satisfactory, the globules present were rounded in shape, but when they deteriorated and became elongated, the end product was lactic acid fermentation. There are four general requirements for such work, as suggested by the papers published by Pasteur on lactic acid fermentation. The conditions to be fulfilled include the following: (1) Optimum conditions must be prevalent in order to study the process of fermentation, (2) the substances used should be from the simplest possible source, (3) the appearance of the organism during the process of fermentation should remain constant, which should be confirmed by careful examination under the microscope, and (4) the characteristic fermentation can be produced by even a minute trace of the presumptive cause. In 1857, Pasteur published the first paper on alcoholic fermentation. In accordance with the catalytic theory of von Liebig and Berzelius, the ferment takes nothing from the fermentable material and gives up nothing during the fermentation process. However, when the ingredients were weighed prior to the commencement of, and following, fermentation, it was clear that the yeast cells were taking something from the sugar. The sugar's breakdown into carbon dioxide and alcohol was associated by Pasteur to the living processes, wherein the part of the material for the yeast was provided by sugar. The role of yeast in alcoholic fermentation was affirmed by Pasteur in 1860. In sharp contrast to von Liebig's assumptions, ethanol and carbon dioxide constituted only 95% of the products of fermentation of invert

sugar (mixture of glucose and fructose). The rest of the 5% included cellulose, succinic acid, and glycerol. Thus, Pasteur proved that yeast indeed took something from the sugar which was not returned; hence he claimed that fermentation is a physiological process. Also, Pasteur was able to produce a yeast type from a defined medium containing inorganic phosphate and ammonium tartrate along with sugar. There was no existence of any substance that can be putrefied by oxygen and make sugar unstable, as had been proposed by von Liebig. Thus, the theory that the origin of yeast is dependent on the action of oxygen on the fermentable liquid was refuted by Pasteur. Here, Pasteur has also encountered the problem in differentiating between fermentation by intact cells and by enzymic action. There was a long-standing difficulty around the confusion regarding the difference between enzymic action and fermentation, and Pasteur gave special attention to this issue. He had a particular difference of opinion with Marcellin Berthelot (1827–1907), an eminent figure in the French scientific community. In 1860, around the same time when Pasteur published his alcoholic fermentation paper, Berthelot provided an important account of how sucrose is broken down into fructose and glucose by beer yeast. Earlier, Eilhard Mitscherlich had given accounts on the activity of yeast extract to produce a levorotatory sugar from cane sugar. Augustin-Pierre Dubrunfaut further went on to show that glucose and fructose constituted the levorotatory sugar. In order to refute the views of Pasteur, Berthelot in 1860 described the method of invertase (β -fructofuranosidase) isolation that can break down sucrose. Pasteur was of the opinion that it was the succinic acid produced during the fermentation process that was responsible for the breakdown of sucrose. Succinic acid, however, could not invert sucrose at all in conditions similar to that which is prevalent during the process of fermentation, and inversion was only possible in an alkaline medium. During the interval from 1860 to 1880, chemists had a definite change in stance concerning the activities of the enzymes and the process of fermentation. The findings of Berthelot and Pasteur are thought to have provoked these changes in attitudes of the scientific community toward the biological processes and helped immensely in building up and shaping the field of

biochemistry during its periods of infancy. In 1878, Wilhelm F. Kühne, in order to remove the confusion surrounding the double meaning of the term *ferment*, introduced the term *enzyme* to designate the soluble ferments.

Contribution of Eduard Buchner

Eduard Buchner started his work with Adolf von Baeyer in chemistry in 1884 and with C. von Naegeli in botany. His brother Hans was his special supervisor at the Botanic Institute, Munich. In 1885, his first publication explained the effect of oxygen on fermentation. He received special stimulation and assistance for his research in organic chemistry from the likes of H. von Pechmann and T. Curtius. It was possible for him to continue his studies with the aid of the Lamont Scholarship. With grant aids from von Baeyer, Buchner established a small laboratory of chemical fermentation, where he performed his experiments on chemical fermentations. The rupture of the yeast cells was experimentally studied by him in the year 1893, although the board members of the laboratory believed that this would accomplish nothing. In 1907, Buchner went on to receive the Nobel Prize for his discovery of cell-free fermentation and contributions to biochemical research. In 1897, the experimental design was set up with the production of a cell-free extract from yeast cells. Buchner went on to show that the cell-free extract could ferment sugar. The theory of vitalism received a severe blow when it was shown that the process of fermentation did not require the presence of living yeast cells. He prepared the cell-free yeast extract by a combination of kieselguhr (diatomaceous earth) and quartz with dry yeast cells after pulverizing the dry cells with mortar and pestle. With the cell contents coming out of the yeast cells, the mixture became moistened. Following this step, a press was used for the mixture to be passed through that would result in the production of the *press juice* that contained fructose, glucose, or added maltose. Carbon dioxide sometimes was found to develop. It was found through microscopic analysis that the extract contained no living yeast cells. Buchner believed that the proteins secreted by the yeast cells in their environment were responsible for the fermentation of the

sugars, but later it was confirmed that it was within the yeast cells that the fermentation took place.

Sugar Metabolism and Enzyme Specificity

Toward the end of the 19th century, Emil Fischer emerged as a leading organic chemist who today is regarded as the founder of carbohydrate chemistry. The first ideas about enzyme specificity and its molecular mechanisms were drawn from his studies on sugar metabolism in yeast. A new hexose, mannose, was prepared in 1888 by Fischer and Hirschberger, which at room temperature could be fermented avidly by beer yeast. Yeast was earlier encountered by Fischer when his father had invested in a Dortmund brewery. He used sugars such as galactose, mannose, and glucose to test for the fermenting ability of beer yeast. He found that fermentation was possible for the D-sugars only. Thus, he could separate this form from the racemic mixtures with L-forms. Similar approach was undertaken by Pasteur for the tartaric acids. Subsequently, the osazones and hydrazones of the corresponding L-sugars were characterized by Fischer. Prior to this, Fischer was able to discover the reaction between the sugars and phenylhydrazine, which was later used in deriving the configurations of the sugars. The sugars produce characteristic osazone crystals; this was an important finding since previously the major handicap of sugar chemistry was the problem in obtaining sugar crystals. In line with the shift of Pasteur from chemistry to biology, Fischer, after establishing the sugar configurations and monosaccharide classifications, thought of applying his findings to biological research. To save material, he made use of a very small fermentation tube, since the sugar preparations were often laborious and the experiments needed to be repeated frequently. Until exploited by Fischer, Pasteur's finding about the potential ability of the microbes to discriminate between the L- and D-substrates received little attention. Accordingly, Fischer observed that only the D-forms of the sugars like galactose, mannose, and glucose were fermented by the microbes and not their L-counterparts. Additionally, he found that the different sugars were fermented by different yeasts and their fermentation ability also

varied considerably, which can be largely attributed to their structural characteristics. Fischer in 1894 developed the image of lock and key that formed the basis of enzyme-substrate complex formation, as developed later on by Leonor Michaelis, Maud Menten, and Victor Henri. It is also implicated in the idea of substrate destabilization or transition state stabilization and in the role of selective binding energy—the basis for substrate activation theory by J. B. S. Haldane, enzyme transition state complementarity by Linus Pauling, and induced fit theory by Daniel Koshland. All these eminent scientists have acknowledged the findings of Fischer as the basis for their works.

Kühne and Physiological Chemistry

Wilhelm Kühne joined Heidelberg in 1871 to replace Helmholtz. Here in Heidelberg, Kühne went on to produce some of the greatest discoveries of his life, including the digestion of protein and the chemistry of vision. He continued to work selflessly and enthusiastically in Heidelberg and left a rich legacy in the history of biochemistry and physiology. Besides his mainstream work, his research included a series of miscellaneous studies. Kühne published papers on the origin of hippuric acid, artificial diabetes in frogs, resuscitation of carbon monoxide, poisoned dogs, the cause of jaundice, blood ozone content, occurrence of ammonia in blood, cholera transfer and treatment, and on many other aspects of physiological chemistry, thereby demonstrating his breadth of talents and versatility of interests. Living muscle fiber sarcoplasm was considered by Kühne as a protein-rich fluid whereby clotting of the proteins takes place during heat or death rigor. Kühne observed sarcoplasm fluidity with the accidental finding that inside a single frog muscle fiber, a living nematode could pass through the cross striations by performing active movements. This was when the concept of myosin came to the fore. (Albert Szent-Györgyi, a hundred years later, would study the phenomenon of muscle contraction based on Kühne's myosin discovery.) Kühne then turned his focus to the study of digestive enzymes and ultimately went on to discover trypsin, which can be regarded as a landmark discovery in the field of proteomics research. Besides

discovering trypsin, his efforts also identified a number of protein digestion products. Using a mixture of dog pancreas and beef fibrin, he observed that fibrin solubilized with the concomitant appearance of peptone, leucine, and tyrosine. In a piece of dog small intestine which still had in place the pancreatic duct, similar digestion products could be obtained. The digestive function of pancreatic juice was confirmed, and Kühne found that this activity was different from the role, earlier discovered, that gastric juice performs. At the Naturhistorisch-Medizinischen Verein, Heidelberg, Kühne delivered a lecture in 1876. Here, he went on to introduce the term *enzyme* to designate unorganized ferments such as ptyalin and pepsin. The actions of these substances could already be distinguished from that of the living cells known to contain ferments in an organized specialized form. The pancreatic powerful proteolytic enzyme was termed *trypsin* (the cleaver). Kühne is also credited with the extraction of trypsin and examination of some of its properties. This was followed by a series of other observations in relation to trypsin, such as the fistula-mediated collection of pancreatic juice; protein clotting by this juice; observation of the pancreatic cells that secrete this juice (which undergo a change in form during secretion); and the microscopic observation during rest and digestion of the rabbit pancreas. With Ewald, Kühne tested the activity of trypsin on different animal tissues, being aware of the potential usefulness of trypsin in histology. His outstanding understanding of physiology and chemistry allowed him to bring insights from both to solidify the basis of biochemistry. Trypsin was unsuccessful in digesting some of the tissue elements, and these he was able to isolate as keratinized and collagenous structures. This reignited Kühne's earlier interests in protein chemistry and protein digestion, and with Russell Henry Chittenden he resumed his works on these aspects of protein biology. He had earlier established that proteins and carbohydrates were distinct in their composition by employing alkaline and acid hydrolysis to obtain some of the products of protein hydrolysis that were entirely distinct from that of the carbohydrate hydrolysates. The research was extended by Kühne to make it more physiological and relevant from the biological perspective by using pancreatic and

gastric juice enzymes. Two of the general groups of substances found as the early products of digestion, namely peptones and albumoses, had previously been found by different investigators as products of alkaline and acid hydrolysis of proteins. Eventually, a number of amino acids could be detected. It was quite obvious that Kühne and Chittenden were successful in identifying the sequential nature of intestinal (pancreatic) and gastric digestion of proteins, since the breakdown of proteins could proceed to a certain extent in an acid-trypsin environment and required the alkaline-trypsin environment to proceed further. Knowledge regarding the difference in action of trypsin and pepsin on the peptides has been much extended to the present day and terms such as *antipeptones*, *antialbumoses*, *hemipeptones*, and *hemialbumoses* that were used frequently by Kühne and Chittenden are no longer existent. However, it cannot be denied that these pioneering works are the cornerstone for the field of proteolysis, which has immensely enriched the sphere of biochemistry.

Syed Feroj Ahmed

See also Biochemistry, 20th Century

Further Readings

- Dubos, J. (1951). Louis Pasteur: Free Lance of Science, Gollancz. Quoted in Manchester K. L. (1995): Louis Pasteur (1822–1895): Chance and the prepared mind. *Trends in Biotechnology*, 13(12).
- Fischer, E. (1894). Synthesen in der Zuckergruppe II. *Berichte der Deutschen Chemischen Gesellschaft*, 27.
- Kühne, W. (1877). Über das Verhalten verschiedener organisirter und sog. ungeformter Fermente. *Verhandlungen des naturhistorisch-medicinischen Vereins zu Heidelberg. Neue Folge*, 1.
- Mitscherlich, E. (1842). Ueber die chemische Zersetzung und Verbindung mittelst Contactsubstanzen. *Annales des Chimie et des Physique*, 55.
- Pasteur, L. (1858). Memoire sur la fermentation appele'e lactique. *Annales des Chimie et des Physique*, 52.
- Payen, A. & Persoz, J. F. (1833). Mémoire sur la diastase, les principaux produits de ses réactions et leurs applications aux arts industriels (Memoir on diastase, the principal products of its reactions, and their applications to the industrial arts), *Annales de Chimie et de Physique*, 53(2).

Ukrow, R. (2004). Nobelpreisträger Eduard Buchner (1860–1917): Ein Leben für die Chemie der Gärungen und - fast vergessen - für die organische Chemie.

BIOCHEMISTRY, 20TH CENTURY

Biochemistry may be defined as the study of different chemical processes in all living organisms. These processes typically govern all life functions, as for example by controlling information flow through biochemical signaling networks and the intrinsic flow of chemical energy by way of metabolism.

The history of biochemistry spans more than eight centuries. Its origins can be traced to the ancient and medieval societies of China and India, with subsequent advancements made in the Islamic world during the centuries that followed. With the introduction of clinical trials and clinical pharmacology in the modern era, biochemistry had continued to make significant breakthroughs, although it was only in the year 1903 that the term *biochemistry* was proposed by Carl Neuberg, a German chemist, to denote a single interdisciplinary subject combining elements of what hitherto had been two separate disciplines, biology and chemistry. Over the course of the 20th century, biochemistry gradually developed into an essential discipline within the life sciences.

Much research deals with the basic structures and effective functions of cellular components such as carbohydrates, proteins, nucleic acids, lipids, and many other biomolecules, although the study of various natural processes involving close-linked networks rather than those of individual molecules is quickly becoming the focus of biochemistry. During the last half century, biochemistry has succeeded in explaining a variety of life processes, and as a result many disciplines, from medicine to botany, are seriously engaged in biochemical research. Today, it has become the main focus of pure biochemistry to understand how biological molecules actually give rise to the processes that typically occur within living cells, which in turn significantly contributes to the understanding of whole organisms.

Among the vast numbers of many different biomolecules that exist in the living system, most are large, highly complex and molecules (called

biopolymers), which are composed of many similar repeating subunits (called *monomers*). Each of these classes of polymeric biomolecules has a very different set of various subunit types—for example, a *protein*, which is essentially a polymer whose different subunits are selected from a typical set of 20 or more amino acids. *Carbohydrates*, in the form of polysaccharides, are formed by linking monosaccharides or oligosaccharides; *lipids* are typically formed from combining fatty acids and glycerols; and the nucleic acids are formed from the simple nucleotides. Biochemistry also studies the basic chemical properties of the very important biological molecules, which includes proteins with catalytic activities, in particular the hard-core chemistry of the enzyme-catalyzed reactions. Furthermore, the metabolism of specific cell types and the working of the endocrine system have also been quite extensively described. Some other areas of focus in biochemistry include the understanding of the genetic code as embodied in the nucleic acids deoxyribonucleic acid (DNA) and ribonucleic acid (RNA); cell membrane transport; protein synthesis; and the phenomenon of signal transduction, by which a chemical or physical signal is transmitted through a cell as a series of molecular events. Thus, biochemistry as a field of study is highly dynamic and has the potential to continue to develop with time. The following sections discuss some of the significant discoveries and achievements in biochemistry during that period and identify the key subfields and related disciplines that have emerged since the turn of the 20th century.

Major Developments in the 20th Century

Before the turn of the 20th century, the notion was widely held among scientists that living cells were occupied by a substance called *protoplasm*, in whose huge molecules obscure and perhaps incomprehensible chemical changes continuously took place. This view would become discredited when it was discovered that complex cell constituents like the proteins and enzymes could be effectively isolated and carefully crystallized. In 1901, for example, the English biochemist Frederick Gowland Hopkins, with his graduate student Sydney W. Cole, in 1901, discovered the amino acid tryptophan to be a typical constituent of many proteins.

Enzymes

In 1897, the German chemist Eduard Buchner began to study the specific ability of yeast extracts to efficiently ferment sugar despite the absence of any living yeast cells. In a series of critical experiments at his laboratory in Berlin, he found that fermentation of sugar occurred without the presence of living yeast cells owing to the action of some chemical compound in the extract or mixture. He determined this to be an *enzyme* (i.e., a catalytic protein that brought about the fermentation of sucrose), which he named *zymase*. In 1907, Buchner received the prestigious Nobel Prize in Chemistry for his outstanding biochemical research that led to the discovery of typical cell-free fermentation. Following Buchner's example, researchers started naming enzymes according to the kind of reaction they perform in the system. For example, *lactase* is the name of the enzyme that cleaves lactose; *polymerase* is the enzyme that catalyzes the formation of DNA polymers. Many early workers noted that enzymatic activity was typically associated with proteins, but many scientists at that time, including the 1915 Nobel laureate Richard Willstätter, vigorously argued that the proteins were merely the carriers for true enzymes and that these proteins were virtually incapable of performing catalysis. However, in 1926, James B. Sumner clearly showed that urease, a common enzyme, was basically a pure protein and efficiently crystallized it; he again did likewise for another enzyme, catalase, in 1937. Further, the distinct conclusion that many pure proteins can act as enzymes was definitively proved by John Howard Northrop and Wendell Meredith Stanley, who crystallized the three digestive enzymes trypsin, pepsin, and chymotrypsin. For their pioneering efforts Sumner, Northrop, and Stanley were jointly awarded the Nobel Prize in Chemistry in 1946. The critical discovery that enzymes could be effectively crystallized meant that scientists could eventually solve their molecular structures by the technique of X-ray crystallography. This was for the first time done for the enzyme *lysozyme*, which is found significantly in saliva, tears, and egg white, and which is capable of breaking down the cell wall of many gram-positive bacteria. This structure was actually solved by a group of researchers led by David Chilton Phillips, and the

work was published in the year 1965. This crystalized high-resolution structure of lysozyme marked the beginning of the field of structural biology and also led to the effort to understand how these enzymes actually work at the atomic level.

Metabolism

One of the most significant and prolific of 20th-century investigators was the biochemist and physician Hans Adolf Krebs (1900–1981), who made major contributions to the study of metabolism. He himself discovered the urea cycle and later, while working with Hans Kornberg (1928–2019), discovered the glyoxylate cycle and the TCA or citric acid cycle. These significant discoveries led to Krebs being awarded the Nobel Prize in Physiology in 1953, which he shared with another well-known biochemist, Fritz Albert Lipmann, who was in addition the codiscoverer of the essential factor *coenzyme A*. In 1960, the biochemist Robert K. Crane revealed his extraordinary discovery of the phenomenon of sodium-glucose cotransporters as the major mechanism for glucose absorption in the intestine. This was the first significant proposal of a typical coupling between the basic fluxes of a substrate and an ion, and it has been viewed as having efficiently sparked a new revolution in biology. This discovery would not have been possible, however, without the work of Emil Fischer, who had actually discovered the structure and chemical properties of the molecule glucose, for which he was awarded the 1902 Nobel Prize in Chemistry, nearly 60 years before. Because the process of metabolism entails the breakdown of larger molecules into simpler ones (catabolic processes) and the synthesis of larger complex molecules from these simple particles (anabolic processes), the importance of glucose and its involvement in the formation of adenosine triphosphate (ATP) are fundamental to our understanding of metabolism. Glycolysis, also known as the *Embden–Meyerhof–Parnas (EMP) pathway*, named after the three scientists who discovered it, is one of the main pathways by which breakdown of glucose occurs under both aerobic and anaerobic conditions. These three researchers discovered that the glycolysis is actually a strongly determinant process, which is essential for the functional efficiency of the human body.

Understanding this pathway led to the opening of a new field of clinical biochemistry that participates in the study of metabolic disorders. Subsequently, it was also possible to delineate the difference between the glycolytic pathways in the mammalian and microbial systems.

Major Instrumental Advances

Since the mid-20th century, biochemistry has significantly advanced with the rapid development of more and more new techniques such as electron microscopy, chromatography, radioisotopic labeling, protein nuclear magnetic resonance spectroscopy, X-ray diffraction, and studies of molecular dynamics simulations. These extraordinary techniques have enabled the discovery and detailed analysis of many significant cellular functions, such as the citric acid cycle and glycolysis.

Modern biochemistry was revolutionized after the invention of the primary gene amplification technique called the *polymerase chain reaction* (PCR), which was developed by Kary Mullis in the 1980s. This technique basically allows copying a single gene, which is then gradually amplified into millions of copies and has therefore become a cornerstone in the standard protocol for any biochemist who wishes to work with gene expression. PCR is used not only for basic gene expression research but also in aiding research and clinical laboratories in diagnosing leukemia, lymphomas, and other malignant diseases. Without PCR development, many important advances in the field of protein expression study and bacterial study would not have come to fruition. Along with the development of the basic theory and the process of PCR, the invention of the instrument called the *thermal cycler* proved to be critically important, since the use of PCR would not have been possible without it. This is a clear instance of the combined value of the advancement of technology and the painstaking research that resulted in the development of key theoretical concepts in biochemistry and their subsequent application.

Vitamins

In 1912, the Polish biochemist Casimir Funk put forward the arresting theory that scurvy, beriberi, rickets, and perhaps pellagra were actually

caused by the lack or deficiency of some special substances in the diet, which he named *vitamins*. With the typical exception of vitamin B₁₂ and folic acid, by 1940 all of these substances had been mostly identified, purposely isolated, and chemically characterized.

Before the onset of World War II, not only the structures of almost all the vitamins had been discovered but their actual modes of action had also been largely recognized. Vitamin B₁ (thiamine) deficiency was found to cause polyneuritis in the kidney due to the faulty utilization of pyruvate. It was actually the cofactor thiamine pyrophosphate which proved to be extensively needed for the typical processes of natural or biological decarboxylation of molecules like pyruvate and others. During this era, nicotinic acid was found to be an important growth factor. In addition, its presence is required for the effective biosynthesis of nicotinamide coenzymes, which play an important and fundamental role as typical hydrogen carriers, in both aerobic and anaerobic organisms. This latter fact makes it quite possible that nicotinamide adenine dinucleotide (NAD) was among the first substances ever formed in living cells, perhaps millions of years ago when life on earth greatly flourished in an anaerobic atmosphere. Subsequent observations pointed out the efficacy of pyridoxine, or vitamin B₆, as another important cofactor, shown later to play an important role, in the form of its pyrophosphate, in the amino acid transaminations and decarboxylation.

By midcentury, most of the now well-known vitamins had been discovered and their typical modes of action explained. Pantothenic acid was actually found to be an integral part of the coenzyme A molecule. At this time, biotin had been discovered, and its structure had been totally elucidated, then its mode of action in the path of causing major carbon dioxide assimilation in animal cell metabolism was largely clarified by Feodor Lynen and colleagues.

Molecular Genetics

With its roots in biochemistry and closely allied with molecular biology, another discipline of life sciences with vast potential is molecular genetics. A prime target of its investigations is the genetics of enzyme-deficiency diseases and their

basic clinical aspects. Molecular genetics offers tools for anticipating major health risks and for the prevention of diseases. Yet another related field of importance is molecular pharmacology, which aims to alleviate and cure our aches, pains, and diseases.

Neurochemistry

Another important domain, which emerged in the latter half of the 20th century, is a neurochemistry, which studies the chemistry of the nerves and the brain. Here, the cell surfaces play a fundamental role, as do the *synapses*, the gaps between the terminal dendrites of adjacent neurons where electrochemical events occur between communicating cells. At each synapse there is a ceaseless and rapid surge of potassium, sodium, calcium ions as well as some other ions and amino acids, various types of organic substances, and at specific gates or pores; that is, points having special permeability. A neuron can effectively pass an electrochemical signal to another by a chemical messenger through the synapse and a very long, quite sustained, chained effect is accomplished whereby a typical stimulus at a particular sensory organ is finally communicated to the brain and a response by the brain is generated. The vesicles that exist in the synaptic cleft of the neuron possess a high concentration of chemical messengers. This organized phenomenon of typical exocytosis is the foundation of our understanding of perception and of life functions ranging from simple reflexes to conscious awareness and abstract thought.

Conclusion

Biochemistry interfaces with both biology and chemistry and is concerned with the various chemical processes that occur within living cells. Contemporary biochemistry emerged from what in the late 19th and early 20th century was called *physiological chemistry*, which dealt chiefly with extracellular bodily fluids and the processes of digestion. Biochemistry as such is thus largely, though not quite exclusively, a 20th-century discipline. Molecular biology, which came into existence as an offshoot of biochemistry itself, has come to mean a study of the three-dimensional

structure and function of such biologically relevant and important biomolecules as nucleic acids and proteins. Molecular biology is basically as much a typical interface of biology with chemistry as of general biology with physics. In many such respects biochemistry and molecular biology represent the basic realization of the dream of the mechanistic biologists of the early 20th century, who were convinced that the most fundamental biological processes could of course ultimately be well understood in terms of the fundamental laws of both physics and chemistry.

Syed Feroj Ahmed

See also Biochemistry, 19th Century

Further Readings

- Bartlett, J. M. S., & Stirling, D. (2003). A short history of the polymerase chain reaction. *Methods in Molecular Biology*, 226.
- Florkin, M., & Stotz, E. H. (1970–1995). *Comprehensive biochemistry, section VI (Volumes 30–39): A history of biochemistry*. Elsevier.
- Fruton, J. S. (1999). *Proteins, enzymes, genes: The interplay of chemistry and biology*. Yale University Press.
- Funk, C. (1912). The etiology of the deficiency diseases. Beri-beri, polyneuritis in birds, epidemic dropsy, scurvy, experimental scurvy in animals, infantile scurvy, ship beri-beri, pellagra. *Journal of State Medicine*, 20.
- Hopkins, F. G. (1913). The dynamic side of biochemistry. *Report of the British Association for the Advancement of Science*, 652.
- Kohler, R. E. (1973). The enzyme theory and the origin of biochemistry. *Isis*, 64(222).
- Quastel, J. H. (1985). The development of biochemistry in the 20th century. *Molecular and Cellular Biochemistry*, 69(1).
- Simoni, R. D. (2002). The discovery of Glutathione by F. Gowland Hopkins and the beginning of biochemistry at Cambridge University. *The Journal of Biological Chemistry*, 277.
- Sumner, J. B. (1926). The isolation and crystallization of the enzyme urease. Preliminary paper. *Journal of Biological Chemistry*, 69.
- Templeton, N. S. (1992). The polymerase chain reaction. History, methods, and applications. *Diagnostic Molecular Pathology: The American Journal of Surgical Pathology, Part B*, 1(1).

- Weckbecker, A., Gröger, H., & Hummel, W. (2010). Regeneration of nicotinamide coenzymes: Principles and applications for the synthesis of chiral compounds. *Advances in Biochemical Engineering/Biotechnology*, 120.
- Wilson, B. A., Schisler, J. C., & Willis, M. S. (2000). Sir Hans Adolf Krebs: Architect of metabolic cycles. *Laboratory Medicine*, 41(6).

BIOCHEMISTRY, CONTEMPORARY

Biochemistry is a discipline that examines and implements new knowledge of the physical, biological, and chemical processes within organisms, how these constitute interrelated systems, and how their structures determine their physiological functions. The study of biochemistry describes the structures and mechanisms of life in molecular terms and the evolution of complex physiological functions. The chemical nature of simple organic molecules yields a world of its own—one in which 20 standard amino acids assemble to generate essential functions of complex proteins and structures. This examination of contemporary biochemistry aims to discuss the physical and chemical processes involved within organisms and how larger systems ultimately develop. After a presentation of some basic concepts, the entry provides an explanation of the features of cells, followed by a discussion of the role of nucleic acids in living systems. Next, a detailed review of protein structure and functions precedes a discussion of drug development and technology, which identifies the primary targets of drug research. The entry continues with an account of cellular metabolism and disease and concludes with an exploration of the biochemistry of the brain.

Fundamental Concepts

Life consists of three major properties: (1) the ability to transform energy and matter, (2) the ability to grow and reproduce, and (3) the ability to respond to and modify the environment. Traditional undergraduate biochemistry curriculums introduce the chemistry of life sequentially, separating general chemistry from biochemistry by a year's study of organic chemistry. While governed

by the physical and chemical laws of the nonliving universe, the distinction between the subjects lies in differentiating living organisms from nonliving matter. For example, carbon dioxide is an organic molecule and is also a biochemical substrate for the enzyme carbonic anhydrase, which catalyzes the production of bicarbonate and protons. The resulting Bohr effect is a biological phenomenon in which carbon dioxide reduces hemoglobin's affinity for oxygen. Carbonic anhydrase is itself a biomolecule, yet more complex and animated in its function than carbon dioxide.

Biomolecules are a collection of organic molecules, often complex chemical structures, which, in an organized and hierarchical manner, maintain and perpetuate life governed by the many physicochemical principles associated with the movement of energy through matter. Yet, when trying to understand the chemical processes within and between organisms, what is already known often needs refinement. Acquiring new knowledge of how these molecules interact allows biochemists to formulate potentially alternative models on established concepts. Contemporary biochemistry builds upon traditional knowledge to accumulate a new understanding of perceived biological phenomena that often lead to novel approaches in biotechnology.

Additionally, *biomedicine* is the application of biochemical theories toward insightful discoveries in developmental biology, anatomy, and physiology. Contemporary scientists use fundamental core concepts to find key ways to tackle new problems, so as to develop an increasingly comprehensive understanding of human development and disease, with a focus on case studies and clinical application. Concepts that constitute contemporary biochemistry and distinguish them from traditional biochemistry aim to answer many questions about science and life, and how the two together can broaden perspectives about cellular interactions.

The Cell

All cells share similar chemical features. They contain approximately 70 percent water by mass. They have lipid membranes that separate and protect the intracellular components from the external environment. The organelles and other significant

cell components are composed of four major macromolecules: *proteins*, *lipids*, *carbohydrates*, and *nucleic acids*. Simple organic molecules are used as building blocks for the larger complex macromolecules. The origins of the cell are not well understood, yet there is striking consistency in the biochemistry of life from the simplest prokaryote to complex organisms. Norman Pace presented an insightful perspective on biochemical unity in which he reasons that the phylogenetic tree of life, based on small ribosomal ribonucleic acid (RNA) sequences, indicates that the three domains of simple cellular life—archaea, eukaryotes, and bacteria—have a common evolutionary origin, and that inorganic molecules may have been used to support early metabolic processes. The relative ordering of evolutionary events between the earlier form and the three domains is the subject of an ongoing debate among scientists.

Understanding the nature of cells through exploration, discovery, inquiry, and consideration has led to accelerated advancements in stem cell research. Stem cells are a unique type of cell, which can either remain unspecialized or differentiate into specific cell lineages. Discovering the active biochemical agents that guide decisions made by stem cells provides insight into tissue repair and wound healing treatments.

Nucleic Acids

Genomic information of the cell is stored and expressed by nucleic acids. Nucleic acids are long, linear polymers formed from four types of monomers. Each monomer consists of a sugar, phosphate, and nucleobase. The sequence of the bases yields the information content of nucleic acids. Information storage and active use through metabolic processes ensure fidelity. Optimizing information transfer is provided at the level of genetic replication, which must access and evolve the stored information to be faithful to the environment. Deoxyribonucleic acid (DNA) is the biomolecule responsible for storing information, and as a memory molecule it possesses the collective history of all the environmental factors that led to the survival of our species. The importance of genetic information was noticed after polymeric DNA was inferred by Oswald Avery in 1944 to contain genetic information. In 1953, Rosalind

Franklin made further contributions when she collected X-ray diffraction data on purified DNA, which led to the development of a structural model of the DNA helix by James Watson and Francis Crick. From this era, experimental discoveries from studies of bacteria and bacterial viruses provided evidence that DNA alone functions to store and transmit information.

The structural and biochemical knowledge of DNA gained since the 1970s is now being used in the emerging fields of nanotechnology and biotechnology. The 1970s development of recombinant DNA technology sparked a revolutionary shift in the world of science. Contemporary scientists can now modify DNA sequences in their endogenous contexts and observe the functional organization of the genome at a systems level. Continued study of prokaryotic organisms such as bacteria and an increased understanding of their adaptive strategies against viral infection led to the discovery of the clustered regularly interspaced short palindromic repeats (CRISPR) system by Yoshizumi Ishino in 1987. In 2020, Jennifer Doudna and Emmanuelle Charpentier received the Nobel Prize in Chemistry for developing the system into a genome editing tool in which the CRISPR nuclease cas9 is targeted by a short guide RNA that recognizes the target DNA of choice and integrates exogenous DNA into endogenous sites. Precise modification of genetic building blocks could further enhance drug development processes and medical therapeutics by introducing new drug targets or correcting harmful mutations. Knowledge of DNA modification is imperative for understanding, teaching, or experimenting in contemporary biochemistry and impacts research in bacterial genetics, prokaryotic genetics, prokaryotic gene regulation, developmental biology, and cancer.

Another area of interest is messenger RNA (mRNA) vaccines, which were released under emergency authorization in response to the COVID-19 pandemic in 2020. As noted by science writer Elie Dolgin, in the journal *Nature*, the history of mRNA vaccines is not straightforward. Multiple threads of interwoven discoveries from many scientists contributed to the mosaic of knowledge that led to the repurposing and delivery of mRNA as a vaccine. As it stands, a small segment of mRNA corresponding to an

encoded area of the viral protein is introduced into the cell. Once translated into protein, an immune response is generated in which antibodies are produced and can recognize and mark invading viruses for destruction. Lipid-based nanoparticles are an essential strategy for mRNA delivery, since mRNA is unstable and unable to cross the lipid membrane owing to its negatively charged backbone. The lipid nanoparticles enclose the negatively charged RNA strand, which enables mRNA transport across the lipid membrane and evades detection by RNases once inside the cell. Many details of this process are still not well understood and are an area of interest for biochemists.

Protein Structure and Function

The unpacking of encoded information into proficient proteins occurs at the ribosome in a process called *translation*. Proteins are composed of amino acids linked together by a peptide bond. The distinct ordering of amino acids determines the primary structure of proteins. The diversity of proteins arises from the unique primary sequences of amino acids and their side chains. Proteins fold to create three-dimensional structures with multiple levels of organization. Secondary structures are held together by hydrogen bonding within the backbone of a single polypeptide chain, which collapses into a foldable tertiary structure driven into self-organization and stabilization by the hydrophobic effect. Noncovalent interactions such as hydrogen bonding, ionic bonds, and dispersion forces occur between side chains and help to further stabilize the tertiary structure of a single polypeptide chain. Additional stabilization is provided when the side chains of the amino acid cysteine are oxidized to form disulfide bridges, which are covalent bonds and stronger than noncovalent interactions. Quaternary structures arise from the same folding principles of tertiary structures but occur between multiple polypeptide chains.

Protein folding does not appear to be a random process. A protein with four disulfide bonds can fold in a matter of seconds, even though there is a 1/105 probability of reforming the bonds correctly. The bacteria *Escherichia coli* can make a functional protein in 5 seconds, which includes translation and folding, but if each possible conformation was sampled, it would take 10^{77} years to sample a

reduced number of conformations for a protein containing 100 amino acids (single vibration is 10^{-13} seconds per conformation; 10 different conformations per amino acid). In the hierarchical model of protein folding, hydrogen bonding helps to stabilize local areas of secondary structure, which may narrow the number of possible rearrangements, with distant areas folding together through the hydrophobic effect. This differs from the molten globule model in which the polypeptide chain collapses into a compact globular structure driven primarily by the *hydrophobic effect*, whereby a gain in the stability of the folded protein is due to the release of ordered water molecules from nonpolar side chains of the polypeptide. The brief time that it takes to fold shows that all the information necessary for protein folding is present in the primary sequence. Computational biochemistry builds upon this knowledge by providing insightful discoveries into the mechanism of protein folding with computational software that predicts protein structures based on sequence and structure homology.

Proteins exhibit a multifunctional nature. At a physiological level, proteins are the primary drivers of metabolic processes and the expression of phenotypic function. At the cellular level, proteins serve as enzymes, hormones, or antibodies; additionally, there are structural and membrane-bound proteins. The geometric and chemical complementarity between proteins and their substrates or ligands ensures selectivity, and at a grander scale it enriches the structural and functional diversity of cellular proteins. *Enzymes* are proteins that catalyze reactions in the human body. They cannot change the direction of the reaction but can lower the activation energy needed to speed up a reaction—in other words, how fast the reaction reaches equilibrium. While most enzymes are proteins, in some cases, RNAs are also capable of catalyzing reactions. As biochemical catalysts, enzymes are often regulated by effector molecules and inhibitors. They are often dependent on cofactors that assist in functional group transformations and redox reactions.

Drug Development and Technology

Academics, clinical researchers, and the private sector continue to collaborate in developing specific compounds against targets. G-protein receptors

(GPCRs), the ribosome, and enzymes are known to be amenable targets for drug discovery. The *ribosome* is the site of protein synthesis, translating genetic information into a linear polymer through the condensation of amino acids. Ribosome inhibitors are among the most successful antimicrobial drugs to treat infections and are a central focus in the development of drug design and mechanisms of inhibition.

Nearly all therapeutic drugs are enzyme inhibitors that perform target-specific functions, and so they continue to play a great role in developing commercial antibiotics and instrumentations. The technological discovery of biosensors, motivated by the need to measure blood glucose levels for patients with diabetes, revolutionized gaining control and analysis of specific conditions. The pioneering efforts of researchers in mutating and inhibiting enzymes have been a significant focus in contemporary science, as they allow for a genome-wide study of enzyme evolution. Through genome sequencing, a recent class of enzymes called *pseudoenzymes* were identified. Pseudoenzymes are proteins that have lost enzyme activity owing to mutations and are now areas of target or anti-target in drug development.

Antibodies are also areas of drug target and development. Monoclonal antibodies—molecules produced in vitro that serve as substitute antibodies—are currently used to fight against infection, since the immune system protects against foreign substances by producing antibodies. Antibodies are specific proteins that target certain antigens and mark them for destruction. They bind to their target cells and kick-start a cascade system to eliminate specified foreign substances. Monoclonal antibodies have high specificity for an antigen or epitope. To effectively design target drugs, scientists are searching to identify mRNA/protein levels to determine gene expression abnormalities and whether the presence of specific proteins correlates with disease progression.

Imitating the body's natural mechanisms for evading foreign pathogens has challenges. In the pathogenesis of a disease, several complex pathways are often involved and driven by proteins. Promiscuous proteins, so called, can sometimes bind to many ligands within living cells, becoming problematic in immunotherapy.

Cellular Metabolism and Disease

Proteins and enzymes affect systems at various levels—from genomic sequences to physiological states of an organism or cell. The dynamic nature of cellular metabolism and disease is a growing concern among contemporary scientists. *Metabolism* is broadly defined as a series of reactions within cells that produce or consume energy for vital processes. It can be subdivided into three classes: anabolism, catabolism, and waste disposal. Traditional biochemists have identified core metabolic networks within these classes as responsible for nutrient utilization and energy production in humans and organisms. These include fundamental processes such as glycolysis, respiration, tricarboxylic acid (TCA) and urea cycles (Krebs), and oxidative phosphorylation. These fundamental processes take food such as carbohydrates or fats and oxidize them into carbon dioxide and water. Energy is extracted and converted into adenosine triphosphate (ATP), the universal currency in biological systems, which pays the price for biological functions such as motion, active transport of molecules across the cell membrane, biosynthesis, and signal amplification.

Research in metabolism is beginning to shift the focus on cellular metabolism toward examining metabolic processes that accompany human diseases. For example, Facundo Fernández and colleagues showed that metabolomic profiling of serous ovarian cancer in mouse models allowed visualization of tissue heterogeneity and lipid alterations in tumor cells, thus providing a tool for understanding disease origin and progression. The regulatory state of a cell or tissue is necessary to realize, as it is often driven by signaling pathways and specific transcription factors, which can impose themselves upon the dynamics of its metabolic state. Contemporary scientists also ask whether the reciprocal idea can also be possible: whether the metabolic state imposes itself on the regulatory state of the cell. Understanding the relationship between cellular-level interactions and biological systems will help discover key proteins and how they function.

Biochemistry of the Brain

The brain is the most complex organ in the human body, composed of an intricate web of synaptic

connections and neurochemicals necessary for a normal functioning central nervous system. The brain accounts for more than 20% of total oxygen metabolism and represents the largest source of energy consumption in a human. Oxidative DNA damage is considered a significant cause of neurological diseases in humans. Mitochondria, a type of organelle within cells, produce reactive oxygen species (ROS) as a by-product of normal cell activity. Michelle Watts and associates have shown that overproduction of ROS and disruption of mitochondria function are consistent causes of many neurodegenerative diseases, such as dementia. The regulation of tissue metabolite supply and cellular energy metabolism is essential to maintaining healthy cellular and systemic function, although prognostic and diagnostic approaches need refinement. Daniel Silverman and colleagues acknowledge that assessing correlations between brain metabolism with neurodegenerative diseases is difficult, since there are only a few cases in which patients are monitored over long-term clinical follow-up. Prolonged studies on the effects of brain metabolism on inner-neuronal connections among a larger patient pool will give rise to answering fundamental questions about cognitive health. Such studies eventually may provide insights into the relationship between memory and biological information, revealing how the retention and active use of information passing through the nervous system and mental processes impacts the way life is experienced and defined.

Systems neuroscience aims to understand how complex interactions between networks of neurons give rise to perception, behavior, and neurodegenerative diseases. The catecholamine neurotransmitter dopamine (DA) plays an essential role in behavioral and cognitive functions relating to memory, motivation, learning, and movement. The progressive degeneration of neurons of the substantia nigra causes a depletion of DA in areas of the brain and can lead to disorders such as Parkinson's disease. Richard Meade and associates report that the etiology of Parkinson's disease remains unknown, but the pathogenic hallmark of the disease is associated with protein misfolding and the presence of alpha-synuclein, which is itself poorly defined both in structure and function. According to Marek Cieřlak and colleagues, dopaminergic agonists, in conjunction

with the dopamine precursor L-DOPA, are used to slow the dopaminergic neurons' progressive death. The complex nature of the brain's chemistry and structure continues to perplex scientists, as poor diagnostic tools and the lack of reliable biomarkers complicate the selection of drug targets and candidate molecules. In addition to stem cell treatments, contemporary biochemists are working to produce efficient drug targets that surpass the blood-brain barrier, a highly selective border of endothelial cells.

Many avenues have been explored with drugs on the study of the brain and the biochemical basis for protein misfolding in neurodegenerative diseases. Yet, there remains a vast landscape of unknowns for contemporary biochemists to explore, particularly in how to unify neurochemical processes with recent machine technologies. Furthermore, abstract notions of human existence and reality are fundamentally created by neurochemical processes, yet the reconciliation of these two concepts is far from complete. Their fundamental claims ask the same questions scientists have asked for centuries: How is the external knowledge of the world possible in science? Questions concerning the meaning of life must be accounted for at a molecular level as well as at the universal level. Although yet to be convincing, Freeman Dyson once offered an explanation to reconcile the higher centers of human consciousness and the lower level of atoms: namely, that consciousness is not just a byproduct of neurochemistry but an active agent, forcing atoms and subatomic particles to make choices between one quantum state and another.

The more options, or choices, the higher the probability. This is one way to interpret entropy. Macroscopic properties emerge from microscopic interactions, which are often driven to rearrange molecules in space by entropy. Choice continues to be an interesting theme to consider, whether it concerns the mind, the differentiation of stem cells, or the quantum states of particles. Contemporary scientists have yet to develop a way to explore the notion of intellectual intuition in human life with neurochemical processes. The reality of things in how they are perceived and how this is coextensive with the mind thinking those things may help scientists to understand whether or how intuition is a form of cognition,

and how organisms can receive sensory data and form understanding. Emerging subgenres of contemporary biochemistry—drug-target kinetics, neuroscience, brain–machine interfaces, metabolic disorders, and nutrition—dive into such ideas surrounding the human condition.

Sami Dhungana and Tammy Y. Campbell

See also Biochemistry, 19th Century; Biochemistry, 20th Century; Biophysics, Contemporary; Cell Theory

Further Readings

- Cieślak, M., Komoszyński, M., & Wojtczak, A. (2008). Adenosine A(2A) receptors in Parkinson's disease treatment. *Purinergic Signalling*, 4(4), 305–312.
- Doglin, E. (2021). The tangled history of mRNA vaccines. *Nature*, 597, 318–325.
- Dyson, F. (2004). *Infinite in all directions: Gifford lectures given at Aberdeen, Scotland April–November 1985* (1st Perennial ed.). HarperCollins.
- Grosjean, H. (2009). *Nucleic acids are not boring long polymers of only four types of nucleotides: A guided tour*. Landes Bioscience, 2000–2013. Available at Madame Curie Bioscience Database. <https://www.ncbi.nlm.nih.gov/books/NBK6489>
- Hughes, J. P., Rees, S., Kalindjian, S. B., & Philpott, K. L. (2011). Principles of early drug discovery. *British Journal of Pharmacology*, 162, 1239–1249.
- Meade, R. M., Fairlie, D. P., & Mason, J. M. (2019). Alpha-synuclein structure and Parkinson's disease: Lessons and emerging principles. *Molecular Neurodegeneration*, 14, 1–14.
- Nelson, D. L., & Cox, M. M. (2017). *Lehninger principles of biochemistry* (8th ed.). W. H. Freeman.
- Pace, N. R. (1997). A molecular view of microbial diversity and the biosphere. *Science*, 276, 734–740.
- Sah, S., Ma, X., Botros, A., Gaul, D. A., Yun, S. R., Park, E. Y., Kim, O., Moore, S. G., Kim, J., & Fernández, F. (2022). Space- and time-resolved metabolomics of a high-grade serous ovarian cancer mouse model. *Cancers*, 14(9), 2262. Published online 30 April 2022. <https://doi.org/10.3390/cancers14092262>
- Silverman, D. H. S., Small, G. W., Chang, C. Y., et al. (2001). Positron emission tomography in evaluation of dementia: Regional brain metabolism and long-term outcome. *JAMA*, 286(17), 2120–2127.
- Watts, M. E., Pocock, R., & Claudianos, C. (2018, June 22). Brain energy and oxygen metabolism: Emerging role in normal function and disease. *Frontiers in Molecular Neuroscience*. <https://doi.org/10.3389/fnmol.2018.00216>

BIOINFORMATICS

Paulien Hogeweg and Ben Hesper coined the term *bioinformatics* in 1978 to refer to the study of informatics processes in biotic systems. Bioinformatics is the application of information technology to the field of molecular biology and more generally asking biological questions with a computer. There was a time when biology happened mostly in dissection labs and test tubes and under microscopes. Owing to the development of genomic technologies, biology has been transformed from a science in which the human effort was mainly oriented toward data gathering to a science that generates a huge volume of data. Many scientists today refer to the next wave in bioinformatics as systems biology, a new approach to tackling new and complex biological questions. The range of possible applications of bioinformatics is enormous: molecular biology, clinical medicine, pharmacology, biotechnology, forensic science, anthropology, and many other disciplines.

Use of Sophisticated Analytic Tools

Bioinformatics is about searching biological databases, comparing sequences, and looking at protein structures. Systems biology involves the integration of genomics, proteomics, and bioinformatics information to create a whole-system view of a biological entity.

The origins of bioinformatics derive from the existence of biological databases. The first bioinformatics or biological databases were constructed a few years after the first protein sequences of amino acids became available, resulting from work on insulin in 1956. After the formation of the databases, tools became available to search sequence databases. Since these early efforts, significant advances have been made in automating the collection of sequence information. Rapid innovation in biochemistry and instrumentation has brought us to the point where the entire genomic sequence of several organisms is known. Projects to elucidate more than 100 prokaryotic and eukaryotic genomes are currently under way. The Internet is the virtual laboratory in which genomic research is now conducted.

In the early 1990s, scientists at the European Organization for Nuclear Energy (CERN) invented the World Wide Web (WWW) technology on the Internet (the now ubiquitous computer network developed earlier in the United States). The Web was the platform that solved many problems of maintenance, update, access, and integration of databases in molecular biology. In a way, without the WWW technology, the Human Genome Project would not have been possible.

The information archive within each organism is its genetic material (DNA and RNA). The human genome is only one of the many complete genome sequences known. The ENCODE Project (ENCyclopedia Of DNA Elements) has the ultimate goal of developing methods for comprehensive identification of functional regions of the human genome, including coding and regulatory regions.

Roderic Guigó, one of the main characters in the race for the genome project culmination, said that life begins when the nucleotides are arranged in the sequence of the genome. It is the particular order of nucleotides in this sequence, rather than its physical and chemical properties, that dictate the biological characteristics of living beings.

The human genome sequence is now complete, and it is joined with 18 archaea, 155 bacteria, and over 30 Eukarya, as well as many other organelle and viral sequences that are now known.

Key databases containing this information include the archive of nucleic acid sequences known as the International Nucleotide Sequence Database Collection, maintained by GenBank, based at the U.S. National Center for Biotechnology Information, in Bethesda, MD; the EMBL Nucleotide Sequence Database, or EMBL-Bank, based at the European Bioinformatics Institute, in Hinxton, U.K.; and the Center for Information Biology and DNA Databank of Japan, located at the National Institute of Genetics in Mishima, Japan.

The archive of amino acid sequences of proteins is maintained by The United Protein Database, a merger of the databases SWISS-PROT, The Protein Identification Resource, and Translated EMBL. Three systems use Web technology to facilitate access to genomic information via distributed databases: ENSEMBL in Europe and NCBI and Genome Browser, both in the United States.

The implementation of genetic information occurs through the synthesis of RNA and proteins.

However, not all DNA is expressed in proteins or structural RNA. Most genes contain internal untranslated regions called *introns*. Some regions of the DNA sequence are devoted to control mechanisms, and for a substantial number of regions, we do not yet understand the function. Proteins, in contrast, show a great variety of three-dimensional configurations with diverse structural and functional roles. Originally, bioinformatics concentrated on the study of the genome, but today it also extends to the study of the proteome, involving the patterns of gene expression and the complex networks of regulatory interactions associated with protein functions.

Alignment: Toward Structure and Function

Aligned sequences of nucleotide or amino acid residues are typically represented as rows within a matrix. In the precomputer era, sequences of DNA, RNA nucleotides, or protein amino acids were assembled, analyzed, and compared manually. Later, these problems were solved using algorithms instead. An algorithm is a complete and precise specification of a method for solving a problem. As soon as computers became available, the computational biologist started to enter these manual algorithms into the machines.

In general, this new discipline called *bioinformatics* could be summarized in the term *alignment*. In bioinformatics, a sequence alignment is a way of arranging the sequences of DNA, RNA nucleotides, or protein amino acids to identify regions of similarity that may be a consequence of functional, structural, or evolutionary relationships among the sequences. Sequence—structure—function: This is now a central concept of both molecular biology and bioinformatics.

Protein sequence alignment has become an essential task in modern molecular biology research. A number of alignment techniques have been documented in the literature, and their corresponding tools are made available as both free-ware and commercial software. The two most commonly used programs are the Basic Local Alignment Search Tool (BLAST) and FASTA. BLAST finds regions of local similarity between sequences. The program compares nucleotide or protein sequences to sequence databases and calculates the statistical significance of matches.

BLAST can be used to infer functional and evolutionary relationships between sequences and to help identify members of gene families. BLAST was developed to provide a faster alternative to the earlier FASTA without sacrificing much accuracy.

These programs are an ideal starting point to determine whether a related sequence, or a family of sequences, already exists in a database. The results from these programs will provide evidence of function, utility, and completeness of the gene product. In the 21st century, the sensitivity of sequence searching techniques has been improved by profile-based or motif-based analysis, which uses information derived from multiple sequence alignments to construct and search for sequence patterns.

Núria Pérez-Pérez

See also Informatics

Author's Note: Adapted from Pérez-Pérez, N. (2010). Bioinformatics. In S. H. Priest (Ed.), *Encyclopedia of science and technology communication*. SAGE.

Further Readings

- Claverie, J. M., & Notredame, C. (2003). *Bioinformatics for dummies* (2nd ed.). Wiley.
- Lesk, A. M. (2019). *Introduction to bioinformatics* (5th ed.). Oxford University Press.
- Serra, R. G. (2007). Bioinformàtica [Bioinformatics]. *Les biotecnologies: Treballs de la Societat Catalana de Biologia* [The Biotechnologies: Works of the Catalan Society of Biology], 58, 11–24.

BIOLOGY, EVOLUTIONARY

Evolution is the study of the history of life on earth and how it has changed over time. One of the most famous statements on the subject is Theodosius Dobzhansky's 1973 dictum, "Nothing in biology makes sense except in the light of evolution" (*The American Biology Teacher*, p. 125). As we shall see, it is equally true to note that nothing in evolution makes sense except in the light of biology. Over almost three centuries of modern thought on the topic, theories of evolution have been shaped by the current status of knowledge about biological processes, most especially the nature of heredity and hereditary variation. The major key to

understanding why evolution theories have changed and continue to do so today is to be found in new technical and conceptual advances in the science of inheritance. We are in a revolutionary period in study of that subject because we can now directly read and modify genomic DNA sequences. Consequently, the topic of evolution currently has more complexity and intellectual ferment than the general public realizes.

Setting the Scientific Stage for Evolutionary Thinking

Although in antiquity there were some precursors to ideas about evolution, the dominance of religious thought about divine creation of all living organisms dominated the Christian and Islamic worlds through the Middle Ages and even survives today among fundamentalist groups of both faiths. The scientific basis for modern ideas about evolution originated in the 18th and early 19th centuries around two related subjects: taxonomy and paleontology.

Although he believed in the fixity of species, the main figure in the origins of scientific classification and the idea of biological relatedness was the Swedish scholar Carl Linnaeus (1707–1778). In 1751, Linnaeus published *Philosophia Botanica*, setting out the bases for taxonomic classification into species and classes of species by defining relationships through particular inherited characteristics plants did or did not share.

The systematic study of fossils and their connection to taxonomy originated with the Frenchman Georges Cuvier (1769–1832). Known as the founding father of paleontology, Cuvier studied vertebrate paleontology in the context of comparative anatomy. He extended Linnaean taxonomy by recognizing that multiple classes could be grouped into higher level classifications called *phyla* and by incorporating both fossils and living species into the classification scheme. The inclusion of fossils of no longer living organisms established the complementary realities of species extinction and survival by distinct but recognizably similar species. While the extinction/survival duality has become critical to evolutionary thought, Cuvier too accepted the fixity of species. He was the first scientific catastrophist and interpreted the fossil record as evidence for recurring

cycles of species destructions in global mass extinction events followed by creation of functionally similar species without an intermediate process of biological transformation.

Descent With Modification, Natural Selection, and the Origin of Species

The transition to thinking about the diversification of life as a biological process of acquiring and modifying inherited characteristics, known as Descent with Modification, had its origins in the late 18th and early 19th centuries. The idea that living organisms could undergo hereditary variation and transform themselves to new life forms arose in Britain and France. In Britain, the major figure was Erasmus Darwin (1731–1802), Charles's grandfather. He had translated Linnaeus from Latin into English with two colleagues. In E. Darwin's most important scientific work, *Zoonomia* from 1794, he anticipated the modern theory of evolution:

Would it be too bold to imagine, that in the great length of time, since the earth began to exist, perhaps millions of ages before the commencement of the history of mankind, would it be too bold to imagine, that all warm-blooded animals have arisen from one living filament, which THE GREAT FIRST CAUSE endued with animality, with the power of acquiring new parts, attended with new propensities, directed by irritations, sensations, volitions, and associations; and thus possessing the faculty of continuing to improve by its own inherent activity, and of delivering down those improvements by generation to its posterity, world without end! (p. 453)

Erasmus Darwin also anticipated survival of the fittest in *Zoonomia* mainly when writing about the “three great objects of desire” for every organism: “lust, hunger, and security...the strongest and most active animal should propagate the species, which should thence become improved” (p. 452).

In France, the first comprehensive statement of evolution by “transmutation of species” came from the French self-described student of “biology,” Jean-Baptiste Lamarck (1744–1829), in his 1809 *Philosophie Zoologique*. Lamarck was the first scientist to state explicitly that all species arise

from other species and illustrate a branching “tree of life” for animals (p. 649). Lamarck's solution to explaining the apparent direction of evolution was to postulate an inherent drive toward greater organismal complexity (“le pouvoir de la vie,” the life force) and environmental inputs by change linked to the use and disuse of characters that distinguish one species from another. The life force postulates an inherent goal-oriented drive or teleology to the evolutionary process, different from Erasmus Darwin's “lust, hunger, and security,” and has been denounced by modern evolutionary biologists as unscientific. The second adaptive factor is the well-known “inheritance of acquired characteristics.” That process has been singled out by Darwinists to distinguish Lamarck from Charles Darwin as an evolutionary theorist, although Darwin himself explicitly subscribed to the same idea in *Origin of Species* (chapters 5, 6, etc.).

The turn to what is considered by most people to be the definitive theory of evolution came in 1858 with the publication of papers by Alfred R. Wallace and Charles Darwin entitled jointly “On the Tendency of Species to Form Varieties; and on the Perpetuation of Varieties and Species by Natural Means of Selection” (Vol. 3 of *Zoological Journal of the Linnean Society*). The major advances proposed by Darwin and Wallace were (a) discarding any appeals to unexplained internal forces or drives, (b) the assumption of random variation in inherited organismal properties, and (c) the concept of evolutionary advances resulting from differential reproductive success of organisms with improved adaptive characters in a Malthusian struggle for survival and reproduction. In the Darwin–Wallace theory, the purely objective process of *natural selection* for reproductive advantage determined the course of evolutionary change. Darwin explained this idea in the subtitle to his 1859 book, *On the Origin of Species by Means of Natural Selection, or the Preservation of Favoured Races in the Struggle for Life*.

Since its introduction, evolution “by Means of Natural Selection” has received broad acceptance as a foundational explanation of the evolutionary process. However, there were basic missing elements in Darwinism, as even Wallace came to call it (A. R. Wallace, 1889, *Darwinism*). Understanding “Descent with Modification” requires a theory of hereditary transmission and of how hereditary

variants altering adaptive traits arise. Darwin and Wallace had neither, because the fundamental biology was missing. Moreover, Darwin made some *ad hoc* assumptions about variation, apparently influenced by the uniformitarian views of his Edinburgh geology professor, Charles Lyell (1767–1849). Darwin quoted Linnaeus's statement "*Natura non facit saltus*" (Nature does not make jumps) and insisted on evolutionary advances by gradual change over long periods of time in his first edition: "If it could be demonstrated that any complex organ existed, which could not possibly have been formed by numerous, successive, slight modifications, my theory would absolutely break down. But I can find out no such case" (*Origin of Species*, 1859, chap. 6, p. 189). Only later, in his 1872 6th edition, did Darwin recognize "sports" as discontinuous "variations which seem to us in our ignorance to arise spontaneously. It appears that I formerly underrated the frequency and value of these latter forms of variation, as leading to permanent modifications of structure independently of natural selection" (chap. 15, p. 395). As we shall see later, the issue of punctuated versus continuous change has come to be one of the key issues in debates over evolution theory for the last 80 years.

The Science of Heredity: Mendelian Genetics, the Modern Synthesis and the Central Dogma of Molecular Biology

While Darwin and Wallace were proclaiming the idea of evolution by natural selection, the unknown Czech monk Gregor Mendel (1822–1884) took the first steps toward filling in our knowledge about the biological basis of heredity. Mendel documented the regular inheritance patterns of discrete dominant and recessive factors determining coat color and seed shape in breeding (hybridization) experiments with peas in 1866. His results contradicted the prevailing idea that somehow the properties of the parents were blended in their offspring. Mendel must have appreciated the relevance of his findings to evolution because he sent his paper to Darwin, but Darwin ignored it, as did the rest of the scientific world for over 30 years. In 1900, at least two botanists searching for discontinuous inheritance, Hugo de Vries and Carl Correns, rediscovered Mendel's paper and confirmed his rules governing specific trait inheritance.

The widespread dissemination of Mendel's rules at the very start of the 20th century led to an explosion of research and the development of the science of Mendelian genetics, which initiated the scientific examination of biological heredity. The Danish botanist Wilhelm Johannsen (1857–1927) introduced the term *gen* (gene) in 1903 to denote any kind of inherited factor that follows Mendel's rules and in 1909 coined the terms *genotype* (an individual's complete collection of inherited factors) and *phenotype* (the individual's visible properties influenced by the inherited genes). Genetic and microscopic analysis of model organisms, such as maize (corn) and the fruit fly *Drosophila melanogaster*, studied by Thomas Hunt Morgan and his colleagues, established the chromosomes in the eukaryotic cell nucleus as the physical structures carrying Johannsen's genes from one generation to the next. Microscopic and genetic analysis dissected the chromosomal basis for sexual reproduction by haploid gamete (sperm/pollen, egg/ovule) formation in meiotic cell divisions. These studies defined homologous chromosome pairing as a critical step in meiosis that has two major evolutionary consequences:

1. Meiotic chromosome pairing facilitated recombinational exchanges between homologues to generate novel combinations of gene variants along each chromosome and thereby produced additional hereditary variation. This process was quickly recognized and incorporated into mainstream evolution theory.
2. Pairing also restricted successful mating to organisms which shared the same chromosome and thus contributed to reproductive separation of closely related species that had different "karyotypes" or structural nuclear constitutions. Mainstream evolutionists resisted highlighting this feature of species differentiation.

In the 1920s to 1930s, the mathematically gifted pioneers J. B. S. Haldane, Ronald A. Fisher, and Sewall Wright developed the quantitative discipline of population genetics, which adapted the discontinuous, gene-based approach of Mendelian genetics to the idea of continuous evolutionary change. This new, highly statistical approach treated the distribution of the ensemble of individual gene variants within populations as a

major source of evolutionary variation: Natural selection of individuals with favorable combinations of alleles would change the “gene pool” and provide significant adaptive advantages. The assertion that “evolution is a change in gene frequencies” was made on more than one occasion. Together with geneticists, naturalists, and less quantitative theorists, the population geneticists promulgated the modern synthesis (ModSyn) of Darwinism and Mendelian genetics in the early 1940s as a comprehensive and, they believed, definitive scientific statement of evolutionary change. This quickly became the mainstream view of evolutionary processes in the 1940s and has remained so in the public awareness over the past 80 years.

The ModSyn retained Darwin’s view of random mutations making small differences in phenotypic traits that could be selected over time to produce major adaptive novelties. To these random mutations, the ModSyn added genotypic and phenotypic variability resulting from meiotic recombination, patterns of chromosome transmission, and population-level events. Exceptional occurrences of major genotypic change associated with formation of new species or taxonomic groups were explicitly excluded. In the words of Ernst Mayr, a major figure in the ModSyn: “The proponents of the synthetic theory maintain that all evolution is due to the accumulation of small genetic changes, guided by natural selection, and that transspecific evolution is nothing but an extrapolation and magnification of the events that take place within populations and species” (1970, p. 351).

The early years of biochemical genetics and molecular biology appeared to reinforce the ModSyn. One major unanswered question was, how does genetic information function to specify phenotypic characters? The answer came in 1941 from genetic experiments with the fungus *Neurospora crassa* when George Beadle and Edward Tatum reported that mutants deficient in biosynthetic capacity lacked specific enzymes. Their “one gene–one enzyme” hypothesis was rapidly generalized to “one gene–one protein,” and the information content of the genotype was interpreted to consist of the instructions for the synthesis of an organism’s complete repertoire of proteins, the molecules with the biochemical abilities needed to construct a phenotype.

The 1944 discovery by Oswald Avery and colleagues that DNA carries genetic information in bacteria and by Alfred Hershey and Martha Chase in 1952 that bacterial viruses inject DNA into their hosts when they reproduce quickly led to working out the DNA double helix structure in 1953 by Rosalind Franklin and Ray Gosling, James Watson, and Francis Crick. The structure of DNA immediately answered two questions about this central molecule of heredity: How it is able to duplicate accurately, as required for genotype maintenance during cell division and growth? and How it could specify the sequence of amino acids in protein molecules by a nucleotide sequence code? Working out the details of protein synthesis and the genetic code for amino acids in the products inspired Crick to articulate the “central dogma” of molecular biology, which states that sequence information flows unidirectionally from DNA to RNA to proteins, sometimes from RNA to DNA, but never from proteins to DNA. This tidy division of labor in cellular informatics also fit very well with the ModSyn view of the genotype as the controlling blueprint for organismal characteristics. It also provided a molecular explanation of random genetic change as DNA replication errors.

Diverging Views Prior to the Modern Synthesis

The most arbitrary feature of the ModSyn—and of Darwin’s original theory in *Origin of Species*—was the *ad hoc* assumption that the mutational events generating hereditary variation in the first place must be random accidents of small phenotypic consequence. There was no evidence that this was actually correct. The gradualist assumption was made to give predominance to natural selection as a directing force and it effectively assigned a purely passive role to the organism in the evolutionary process. In discussions of evolutionary change over time, there had always been an ongoing debate between gradualists like Darwin and saltationists, such as Hugo de Vries, who believed that nature does make discrete mutational leaps, which can explain the origins of many diverse life forms. A very different disagreement with evolutionary gradualism came from Russian and American scientists who studied cell biology and

proposed that cell fusions between distinct organisms had generated completely novel hybrids by an abrupt process called *symbiogenesis*, whereby one cell becomes an internalized organelle of the new hybrid. This argument was revived in the 1970s by Lynn Margulis (1938–2011) to explain the origins of eukaryotic cells. Both saltatory and symbiogenetic views were treated as far-fetched fantasies by ModSyn followers.

The most serious contemporaneous intellectual challenge to the ModSyn was posed by the developmental geneticist Richard Goldschmidt in his 1940 book, *The Material Basis of Evolution*. Goldschmidt argued that a fundamental difference exists between the accumulation of localized mutations in particular genes that account for *microevolution* improvements that remain within each species and the occurrence of structural changes to the chromosomes that produce *macroevolution* changes forming new species. Goldschmidt also argued that mutational alterations in embryonic development could produce “hopeful monsters,” organisms with major morphogenetic changes (i.e., “monsters”) that occasionally would provide adaptive benefits (i.e., “hopeful”), as a way to explain major changes in organismal body plans. The mainstream ModSyn evolutionist community dealt with Goldschmidt’s well-documented dissent by a form of scientific excommunication, either totally ignoring him or relegating his views on evolution to what they saw as the clearly preposterous hypothesis of “hopeful monsters.” (Today, of course, the hopeful monster idea is seen as the beginning of a major field studying the evolution of multicellular development, widely known as *evo-devo*.)

Problems Arising After the Modern Synthesis

One major problem came out of the research of a plant geneticist credited with being one of the ModSyn founders, G. Ledyard Stebbins. Stebbins studied the role of hybridization between different plant species to generate new organisms for agriculture. He documented the rapid formation of new plant species with new chromosome constitutions following interspecific hybridization and called this process *cataclysmic evolution* (*Scientific American*, April 1951, p. 54). While

Stebbins himself considered his observations a special exception to gradual speciation, they exposed a major evidentiary problem for ModSyn arguments. No amount of gradual selection for changes in individual traits ever produced a novel species, but a very different process has been generating them in agriculture for thousands of years (at least since the origins of flour wheat in the Fertile Crescent about 10,000 years ago in Stebbins’s 1951 narrative).

The most transformational results challenging the assumptions of the ModSyn came from two markedly different sources in the middle of the 20th century. The first source was Barbara McClintock’s studies of chromosome breakage and repair in maize. McClintock found in the 1930s that what were thought to be X-ray-induced “gene mutations” were actually the products of chromosome breakage repaired by the maize plants. She then set out to study chromosome breakage and repair. She documented processes of major chromosome restructuring and unexpectedly discovered genetically mobile “controlling elements” that could insert at new genome locations, where they reprogrammed patterns of expression by adjacent genes. McClintock’s results received a hostile reception when she first presented them to a scientific meeting in 1952, but she was ultimately vindicated and received a Nobel Prize in 1983, recognizing that her controlling elements are universal in biology.

The second transformational episode was humanity’s worldwide evolution experiment combating bacterial pathogens with antibiotics after World War II. When the ModSyn was formulated, it was still debated whether prokaryotic cells without a nucleus had any genetics at all. But postwar experiments pioneered by medical microbiologists and physicists-turned-molecular-biologists established the science of bacterial genetics. However, the rules were strikingly different from those of eukaryotic Mendelian genetics. Bacteria had so-called *infectious heredity* and could readily transfer DNA from cell to cell, often between unrelated types of cells. When antibiotic resistance emerged as a serious clinical problem in the 1950s and 1960s, it did not originate by localized mutations as had been predicted—and experimentally confirmed in laboratory experiments. The resistance determinants

were carried by transmissible DNA molecules, and these “R-factors” accumulated resistances to multiple antibiotics by acquiring mobile DNA elements analogous to those discovered by McClintock. Bacterial versatility in engineering and spreading novel DNA structures is one of the reasons that multidrug-resistant “superbugs” now threaten our health.

Bacterial genetics was also central to pioneering studies of how cells control genome expression. Such studies began to dissect the “gene” and identified two features of genomes absent from the ModSyn: (1) the presence of regulatory DNA sequence elements needed to control the expression of every genetic locus and (2) the shared use of repeated regulatory elements at multiple loci to establish control networks across the genome. In bacteria, but even more so in complex eukaryotes, these multi-locus networks regulate the expression of protein-coding genes under changing conditions. They are the reason so many phenotypes have complex multifactorial genetic determination. Explaining how multisite networks evolved is a serious challenge for ModSyn gradualism but is a straightforward application of the genome-wide distribution of mobile controlling elements.

What Does Genomics Tell Us?

Now that we are able to read DNA and RNA sequences, we need to formulate a more genomics-based view of evolution. One of the first results of sequencing was the discovery by Carl Woese and colleagues as recently as 1977 that bacteria are not the only prokaryotic organisms on earth. There is an equally ancient, abundant and diverse, but evolutionarily distinct kingdom named *Archaea*. This unexpected discovery ultimately led to molecular confirmation of the symbiogenetic origin about 1.8 billion years ago of the ancestor of all eukaryotes by fusion of a *Proteobacterium* ancestor to the mitochondrion and an *Asgard* archaeon related to the eukaryotic cytoplasm. The eukaryotic nucleus contains sequences of both bacterial and archaeal origin, but how it formed as a separate organ of the cell remains unknown. Genome sequencing has also confirmed the symbiogenetic origins of algae, plants, and other photosynthetic eukaryotes. Thus, the deepest evolutionary divisions in biology that

we can explain arose symbiogenetically, not by Darwinian gradualism.

Other features of eukaryotic genome evolution require the formulation of novel theoretical perspectives. The first of these is the discovery of *horizontal DNA transfer* across all phylogenetic boundaries. Examples include bacterial and fungal DNA sequences encoding digestive enzymes in herbivorous invertebrates and sequences encoding eukaryotic regulatory proteins in bacterial pathogens, where they serve to establish the bacterium as an intracellular parasite. Acquiring horizontally transferred proteins that execute specialized functions means that the evolution of key adaptive traits need not be a slow gradual process. The patterns of horizontal transfers are numerous and complex and indicate multiple modes of DNA acquisition. For some researchers in this area, the Tree of Life has transformed into a branching Web of Life.

One of the most important features of complex eukaryotic genomes, the presence of large amounts of repeated DNA sequences, was discovered pre-sequencing in 1968 by a simple biophysical experiment. The presence of abundant DNA repeats in the genome did not accord with the idea of chromosomes carrying unique protein-coding genes like “beads on a string” and demanded an explanation. The response was to call this repetitive DNA “junk,” which accumulated because it had “selfish” properties of replicating faster than unique coding DNA in the genes. Richard Dawkins even elaborated a whole philosophy on the basis of *The Selfish Gene* (1976). Such ideas notwithstanding, sequencing the human and other genomes plus functional studies have revealed important adaptive features of repetitive DNA elements (which outnumber protein-coding sequences in our genome by a factor greater than 30:1):

1. The majority constituents of repetitive DNA are mobile genetic elements distributed throughout the genome.
2. Many of these dispersed mobile elements have been co-opted in evolution as regulatory sites for proper expression of coordinated genome networks encoding complex traits such as pregnancy in mammals.

3. Other dispersed mobile elements serve as part of the DNA templates for the formation of noncoding “ncRNAs,” which encode no proteins but instead fulfill a rapidly growing list of cell regulatory functions, some of which depend upon the repetitive nature of the sequences they contain.
4. A minor fraction of repetitive DNAs consists of tandem repeat arrays that format the chromosomes for proper replication and distribution at cell division. Some arrays also play roles in formatting the genome for epigenetic control of chromatin structure.

In summary, the repetitive DNA story tells us that genomes do more than encode proteins and that we will always find unexpected features, like regulatory ncRNAs, as we delve deeper into the biological basis for evolution. Repetitive DNA elements are typically the most labile genome components in evolution because the majority of the protein-coding sequences are essential for shared functions and conserved by selection, but the regulatory framework involving DNA repeats can change more freely as species diverge.

Mobile DNA elements have been the subject of comprehensive genetic and biochemical analysis. Although these elements come in many forms (e.g., viruses that integrate into cell genomes, *transposons* that migrate within genomes through DNA intermediates, *retrotransposons* that migrate through RNA intermediates), the proteins and nucleic acids involved in genetic mobility have been thoroughly characterized. In particular, the proteins that execute mobile DNA movements in the genome are nonrandom in their actions on the mobile element itself, and many have been found to be targetable for new insertions to particular sites or classes of sites in the genome. Additionally, various ecological triggers and stresses can activate the mobilization of these DNA elements. Thus, for a significant portion of evolutionary mutability, the changes are nonrandom with respect to both genome location and life history events.

While mobile DNA elements sometimes cause large-scale genome rearrangements, it has unexpectedly been the study of somatic evolution in cancer that has significantly expanded our

understanding of major genome restructuring under stress in the 21st century. Whole genome sequencing of various tumor cells revealed highly nonrandom patterns of multiple DNA rearrangements that had occurred in a single chromosome or in a small subset of the 24 human chromosomes. We know from direct observation that these rearrangements often occurred in a single cell division. The concerted rearrangements are triggered by cell division errors and have received colorful names, such as *chromothripsis* (“chromosome shattering”) and *chromoplexy* (“chromosome weaving”). Although cancer might be considered an aberrant condition and irrelevant to organismal evolution, the same nonrandom multisite rearrangement patterns were quickly found to occur in the human germ line and in other animals and in plants. The detailed biology of chromothripsis involves the ability of eukaryotic cells to encase “lagging” chromosomes or fragments left off the mitotic spindle in special subnuclear compartments and to recruit hyper-mutagenic DNA repair enzymes to catalyze the complex multisite rearrangements and generate a level of genome variability that would have been totally inconceivable at the time of the ModSyn.

The Beginnings of a Genomics-Based Theory of Evolution

The study of phenomena like noncoding DNA and ncRNA regulation, the evolutionary utilization of mobile DNA elements to form genomic regulatory networks, and the potential for rapid multisite genome restructuring combine to provide a dramatically different picture of what a genome is and how it can change in evolution. Rather than merely a string of protein-coding genes, the genome has become an intricately formatted and structured database encoding proteins and RNAs, which feed back onto the genome to ensure the correct expression of specific organismal characters. Rather than possessing a read-only ROM memory system that changes gradually by random copying errors, we know that evolving organisms have the biochemical and cellular tools to rapidly restructure their read-write genome databases in macroevolutionary transformations into new species with novel

adaptive traits. Unlike the ModSyn and the central dogma, the genomics-based theory envisions the organism playing an active role in changing its DNA in a manner responsive to its life history and ecology. This means that punctuated equilibrium should be the default mode of macroevolution. The main questions for the future will probably have less to do with biochemical details than with cybernetic questions about how complex DNA changes successfully produce adaptively useful biological novelties.

James A. Shapiro

See also Genetic Drift; Germ Theory; Infectious Disease Studies; Life Sciences, Contemporary; Punctuated Equilibrium

Further Readings

- Bukhari, A. I., Shapiro, J. A., & Adhya, S. L. (Eds.) (1977). *DNA insertion elements, plasmids and episomes*. Cold Spring Harbor Press.
- Darwin, C. (1859). *On the origin of species by means of natural selection, or the preservation of favoured races in the struggle for life*. John Russel.
- Darwin, C. (1872). *The origin of species by means of natural selection, or the preservation of favoured races in the struggle for life* (6th ed.). John Murray.
- Darwin, E. (1794). *Zoonomia; or, the laws of organic life*. J. Johnson.
- Goldschmidt, R. (1982). *The material basis of evolution, reissued (the Silliman Memorial lectures series)*. Yale University Press. (Original work published 1940)
- Gould, S. J. (1983). Punctuated equilibrium and the fossil record. *Science*, 219(4584), 439–440. <https://doi.org/10.1126/science.219.4584.439>
- Gould, S. J. (2002). *The structure of evolutionary theory*. Harvard University Press.
- Heng, H. H. (2019). *Genome chaos: Rethinking genetics, evolution, and molecular medicine*. Academic Press.
- Huxley, J. (1942). *Evolution: The modern synthesis*. Allen & Unwin.
- Margulis, L. (1970). *Origin of eukaryotic cells*. Yale University Press.
- Mayr, E. (1970). *Populations, species and evolution*. Harvard University Press.
- McClintock, B. (1987). *Discovery and characterization of transposable elements: The collected papers of Barbara McClintock*. Garland.
- Sapp, J. (2009). *The new foundations of evolution: On the tree of life*. Oxford University Press.
- Shapiro, J. A. (2011). *Evolution: A view from the 21st century*. FT Press Science.
- Shapiro, J. A. (2021). What can evolutionary biology learn from cancer biology? *Progress in Biophysics and Molecular Biology*, 165, 19–28. <https://doi.org/10.1016/j.pbiomolbio.2021.03.005>

BIOPHYSICS, 19TH CENTURY

While the discoveries of inorganic chemistry were quickly advancing during the end of the Enlightenment period and the Chemical Revolution, many new avenues of scientific inquiry were being born. In the areas of physics and chemistry in particular, new methodologies such as mechanics (engineering and optics) and chemical production (chemical factories) were quickly laying the foundations for invention and chemical knowledge. In fact, the drive to develop industries contributed to the feverish pursuit of scientific knowledge that could be applied in the industrial arena. Among the subspecialties that began to emerge toward the mid-to-late 1800s was biophysics, the application of physics to investigate biological phenomena. Although the actual origin of biophysics is uncertain, the invention of the battery; increased understanding of animal electrophysiology; and developments in electrochemistry, physical chemistry, thermodynamics, X-ray crystallography, and electromagnetic theory all contributed to the field we now call biophysics.

Origins

Scientific inquiry gradually matured through the ages, documented as far back as the ancient Egyptian and Babylonians, though less organized before Aristotle (384–322 B.C.E.) defined the scientific method. In science, perhaps more so than any other school of thought, knowledge builds upon knowledge. A seemingly inconsequential discovery could later inform a different investigation, bringing forth a clarification that would not have been possible otherwise. Biophysics was born from such disjointed discoveries in early experiments that had a narrow focus but later informed or provided a template for a methodology that could be modified to test a new theory. Examples of

these types of experiments were the discovery of bioluminescence in jellyfish and fireflies in studies by Athanasius Kircher, Isaac Newton's foray into the investigation of the nature of muscle movement (1687), and John Walsh's discovery, with Benjamin Franklin, of the electrical discharge from a torpedo fish or electric ray (1773).

Later in the 18th and 19th centuries, scientists became more interested in elucidating the force that produced movement in animals. The scientific method and experimental work concerning the nature of electricity and its role in animal life allowed scientists to ask questions that had not been possible to evaluate before. Most of the experiments conducted during this time were performed by professors of physics who were interested in describing biological phenomena. One such physicist, Abbé Giovanni Beccaria, professor of physics in Turin, performed experiments in which animal muscles received electrical stimulation. An English surgeon and professor of anatomy and physiology in Göttingen, Albrecht von Haller, hypothesized that electrical matter and "animal spirits" were the same.

Electrochemistry

Understanding the electrical component of muscle stimulation was facilitated by the experiments of Italian physiologist Luigi Galvani in the late 1700s. During an experiment, he had skinned a frog leg and was preparing to rub the leg to produce static electricity. His assistant probed the frog leg with a scalpel that unknowingly had a static charge. When the scalpel touched the leg, a spark of static electricity was produced, and the frog leg jumped. At the time, it was believed that electrical current within the animal produced a contraction and that the source was concentrated in the animal's pelvis.

In an effort to clarify the cause of the frog leg's response to the application of static electricity, Alessandro Volta, an Italian contemporary of Galvani, performed several experiments to reproduce the effects witnessed by Galvani. Volta believed that the frog leg only served as an electrical detector to the external electrical source. Meanwhile, those who sided with Galvani's explanation that the electrical source causing contraction was in the animal tissue, several

experiments were conducted that appeared to confirm Galvani's hypothesis. The animal muscle contracted when placed in contact with charged metals. However, Volta performed experiments using his invention of the voltaic pile, so called, a series of alternating silver or copper and zinc discs that had cloth or cardboard soaked in brine placed between the discs and arranged in a circuit, connected to the frog leg. He demonstrated that the contraction was produced by the contact with dissimilar metals, and the frog leg was only behaving as a detector.

Volta's experiments using the voltaic pile resulted in the invention of the early electrical battery. This invention was immediately recognized by the scientific and industrial community for potential uses, which seemed more important than understanding the electrical potential in animals. Study in the area of electrophysiology did not regain popularity until 1827 when the galvanometer, or an instrument for detecting electric current, was built by German electrophysiologist Emil Du Bois-Reymond. This instrument proved able to detect small potential differences across nervous membranes. A new branch of science, *neurophysiology*, developed from the data that were produced by the galvanometer. It is not unusual for physiology and biophysics experimentation to overlap yet remain distinct fields of study.

In addition to neurophysiology, biophysical experiments investigating diffusion gradients and osmotic pressure in living organisms became an important avenue of study. *Osmotic pressure* is the pressure that develops in a solution of salts when separated from pure solvent by a semipermeable membrane. This phenomenon was first described by Abbé J. A. Nollet, a French physicist. The semipermeable membrane was discovered by French scientist René Dutochet in 1828. The understanding of diffusion in biological systems was furthered by botanist W. F. P. Pfeffer, who obtained the first quantitative measurements. Adolf Fick then took the experimental data generated by Pfeffer and merged it with his understanding of the laws governing the flow of heat. He published the first biophysics textbook, *Die medizinische Physik* (Medical Physics) in 1856.

Additional applications of electrophysiology to biophysical experimentation were devised by the German physiologist Hermann von

Helmholtz (1821–1894), who studied both nerve conduction and the physiology of hearing. In 1849, he studied the speed at which an electrical signal was carried along a nerve fiber. Before this measurement, scientists were under the impression that nerve signals traveled faster than anything their available equipment could measure. He utilized the galvanometer developed by Du Bois-Reymond to act as a timing device. Helmholtz attached a mirror to a needle connected to a sciatic nerve of a frog leg, which reflected a light beam across the room to a scale. He was able to detect transmission speeds between 24.6 and 38.4 meters per second. This information enabled scientists to begin to develop better technologies that could capture transmission speeds of nerve conduction.

Developments in Physical Chemistry

Important to the later applications of molecular biophysics (most commonly practiced today) are the atomic theory, the first and second laws of thermodynamics, and advancement in reaction kinetics, which allowed scientists to describe the structural and functional properties of enzymes and proteins. In the early 1800s, the English physicist John Dalton proposed his atomic theory, which stated that matter was composed of discrete, indivisible atoms. It was later corroborated by Amedeo Avogadro in 1811, by making the distinction between atoms and molecules of a specific substance. At the same time, advancements in understanding about heat, energy, work, and temperature were made. The first law of thermodynamics, which states that heat and work are interchangeable, was proposed by the German physicist Julius Robert von Mayer in 1842. Almost 10 years later, the second law of thermodynamics was presented by the German mathematical physicist Rudolf Julius Emanuel Clausius and William Thomson, saying that when energy is transferred or changes from one form to another in a closed system, disorder (entropy) increases.

Chemical kinetics studies were bolstered by the developments made in the area of gas properties and laws. Between 1860 and 1875, Austrian physicist Ludwig Boltzmann and British physicist James Clerk Maxwell demonstrated that the ideal

gas law could be explained by the kinetic theory of matter. Subsequent analysis in chemical kinetics and the formulation of the laws of chemical equilibrium (e.g., Le Chatelier's principle) were based on the kinetic theory of matter. It was determined that the main factors that influence a reaction rate are the concentration, physical state, and temperature of the reactants, and whether there is a catalyst present in the reaction.

In 1864, Peter Waage and Cato Guldberg formulated the law of mass action, which states that the speed of a chemical reaction is proportional to the amount of the reactants. Also, it was found that in consecutive reactions, the kinetics are derived by the rate-determining step. This step represents the slowest step in the series of reactions and can be identified by the rate equation. Knowing the rate-determining step helps to optimize chemical reactions such as catalysis and combustion. This information was important to being able to test the thermodynamics of enzymes and proteins, which enable biophysicists and biochemists to characterize the function of both.

Advancements in optics were also important to the development and characterization of macromolecules, such as proteins that are crystallized (though this application was not practiced until the 20th century). In 1815, a French polymath named Jean-Baptiste Biot studied plane-polarized light (polarized light results from an electromagnetic wave being filtered so that it oscillates in just one plane of vibration rather than in several, unlike nonpolarized light such as light from the sun or from a candle flame); he focused polarized light on organic substances and noticed that the light could be rotated clockwise or counterclockwise, depending on the optical axis of the molecule. This knowledge contributed to determining the structure of proteins and enzymes by X-ray crystallography, which is a very important role that biophysics has played in the 20th and 21st centuries.

Late 19th Century

An increase in the understanding of thermodynamics and molecular structure helped to pave the way for future experiments in structural biology, which draws upon many physical and biological

disciplines to determine the structure of proteins. Scientists were particularly interested in studying the life-sustaining chemical reactions that happened in living cells. Among the important developments that helped to better understand molecular structures were stereochemistry and thermodynamics, as well as advancements in chemical theories.

In 1874, Dutch physicist Jacobus Hendricus van't Hoff established the foundation of stereochemistry with his experiments on optically active carbon compounds and was able to describe three-dimensional and asymmetrical molecular structures. *Stereochemistry* is the study of the spatial arrangement of atoms in a molecule and detection of a change in that arrangement when the molecule is manipulated (e.g. heated). A short time later, he began to apply the thermodynamics of chemical reactions, leading to defining the order of reactions. *Thermodynamics* is a component of physical chemistry that seeks to describe the physical properties of macroscopic systems—for example, enzymes and proteins. These physical properties consist of sensitivities to temperature, pressure, and volume, which are the classic parameters that are measured. In addition to these parameters, variables such as density, specific heat, compressibility, and the coefficient of thermal expansion can be evaluated to determine the full spectrum of physical characteristics and how a substance behaves in its environment.

Other scientists, such as Comte Claude Louis Berthollet, studied the rate and reversibility of reactions, and Benjamin Thompson, an American physicist, tried to elucidate the mechanical equivalent of heat. In 1875, Josiah Willard Gibbs presented the phase rule, which applies thermodynamics to heterogeneous substances. Swedish chemist Svante August Arrhenius produced the theory of electrolyte dissociation, later called *Arrhenius's theory*, in 1889. In 1906, the German physical chemist Walther Hermann Nernst produced Nernst's theorem, subsequently referred to as *the third law of thermodynamics*, which states that the entropy of a system at absolute zero is a well-defined constant. He also contributed to the advancement of physical chemistry by describing physical properties, molecular structures, and reaction rates of macroscopic molecules.

X-Ray Crystallography

One of the most important scientific modalities for biophysics of the 19th century (and beyond) was X-ray crystallography. Crystallography was invented to analyze crystals and their shapes in the 18th century. In 1784, René Just Haüy discovered that a crystal's face could be described in terms of stacking patterns of blocks of similar shapes and sizes. In 1839, William Hallows Miller was able to assign each facet of the crystal a unique three-integer label. Eventually, it was proven that all crystals are a regular three-dimensional array of atoms and molecules. By the 19th century, a complete catalog of possible symmetries of a crystal were constructed by Johan Hessel, Auguste Bravais, Evgraf Fedorov, Arthur Schönflies, and William Barlow. In the 1880s, a few crystal structures had been elucidated, though the structural information was not completely accepted due to the lack of techniques available to verify it (X-ray).

In 1895, Wilhelm Conrad Röntgen discovered the X-ray, though little was known about this form of electromagnetic radiation. Scientists were uncertain of the characteristics of X-rays but were able to reconcile some of the electromagnetic properties of X-rays according to the wave model of light (Maxwell theory of electromagnetic radiation). The visible wavelength of X-rays was determined by Arnold Sommerfeld to be approximately 1 angstrom. By 1905, Albert Einstein had discovered that X-rays were not only waves of electromagnetic radiation but also photons and exhibited particle-like properties. It was observed that when an X-ray is directed toward a crystal, the electron density in the crystal will cause a scattering of light that is characteristic of the arrangement of the electrons within the structure. Based on these scatter patterns, a structure can be determined mathematically. However, this approach to solving the structure of a molecule did not become prevalent until 1912.

Electromagnetic Theory

Electromagnetic theory is important to the field of biophysics because it provides another avenue to observe crystal structures. Up until the late 1800s, electricity and magnetism were considered separate forces. However, when James Clerk

Maxwell published his 1873 *Treatise on Electricity and Magnetism*, he presented the observation that interactions of positive and negative charges seemed to be moderated by a single force. There were four points to his argument: (1) Electric charges attract or repel each other with an inversely proportional force multiplied by the square of the distance between them (opposite charges attract each other); (2) magnetic poles attract or repel based on their charge and occur in pairs; (3) an electric current in a wire generates a circular magnetic field around the wire, and its direction followed that of the current; and (4) a current is created in a loop of a wire when it is pulled away or toward a magnetic field, and follows the direction of the current. This observation helped to advance our understanding of electromagnetic forces, which led to the development of nuclear magnetic resonance (NMR). NMR allows for determination of the structure of solid materials at high resolution. Additionally, elucidation of intermolecular forces generated by the momentum of an electron's movement was made possible. Scientists in the 20th century were able to employ technological advances such as NMR in determining the structure of organic molecules or substances difficult to crystallize, compared with that observed by X-ray crystallography.

Mandy M. McBroom

See also Biochemistry, 19th century; Chemistry; Medicine, Enlightenment; Physics, 19th century; Physics, Quantum Theory.

Further Readings

- Bragg, W. H. (1908). The nature of X-rays. *Nature*, 78(2035), 665.
- Brouwer, D. H. (2008). NMR crystallography of zeolites: Refinement of an NMR-solved, crystal structure using ab initio calculations of ^{29}Si chemical shift tensors. *Journal of the American Chemical Society*, 130(20), 6306–6307. <https://doi.org/10.1021/ja800227f>
- Glaser, R. (2012). *Biophysics: An introduction* (2nd ed). Springer.
- Glynn, I. (2010). *Elegance in science*. Oxford University Press.
- Lewis, G. N., & Randall, M. (1923). *Thermodynamics and the free energy of chemical substances*. McGraw-Hill.
- Macholl S., Borner, F., & Buntkowsky, G. (2004). Revealing the configuration and crystal packing of organic compounds by solid-state NMR spectroscopy: Methoxycarbonylurea, a case study. *Chemistry*, 10(19), 4808–4816. <https://doi.org/10.1002/chem.200400191>
- Patterson, E. C. (1970). *John Dalton and the atomic theory*. Doubleday.

BIOPHYSICS, 20TH CENTURY

Biophysics is a branch of science that integrates several disciplines (physics, chemistry, and bioengineering) to study biological systems. It is applied to the molecular scale, in determining structures of nucleic acids and proteins; to whole organisms, and even ecosystems. One advantage of biophysics is the ability to set up experiments that elicit exact hypotheses and interpretations. The biophysicist is likely to be the first to employ new technologies (e.g., NMR, nuclear magnetic resonance) to apply complex physical theory to the problems posed in biology.

Another advantage is that the application potential of biophysics to the natural world seems endless. A challenge for the student of biophysics, however, is the disparate opinions on what exactly biophysics is. There is scientific equipoise on the actual definition of biophysics, as it spans so many different applications.

However, despite the disagreement among biophysicists on what biophysics really means, there are areas of application that can be used to categorize the contributions that biophysics has made in the 20th century, leading into the 21st century. Astrobiophysics studies the influence of astrophysical phenomena upon life on planet Earth. Medical biophysics uses the advances in biophysics methodologies for diagnostic and therapeutic purposes. Molecular biophysics is the most rapidly evolving version of biophysics; it characterizes biosystems, molecular structures, dynamic behavior of proteins (single molecules up to supramolecular structures), and viruses. Detlev Wulf Bronk, a renowned American physicist and administrator, established

biophysics as a distinct scientific discipline toward the middle of the 20th century.

Over the decades that followed, as technological advances were made, so too did the selection of instrumentation and equipment help to overcome a difficulty in experimentation. Some of the biophysical methods used to study biological systems include X-ray diffraction, chromatography, macromolecular crystallography, proteolysis, NMR spectroscopy, electron paramagnetic resonance, cryo-electron microscopy, multi- and small-angle light scattering, ultrafast laser spectroscopy, dual polarization interferometry, and circular dichroism.

Notable Inventions

One of the most important tools used to determine the structure of a macromolecule is *X-ray crystallography*. By the early 1900s, physicists were beginning to understand the possible scope this technique could have; most certainly, it was the best tool for gathering atomic data on crystals. These data could be mathematically plotted and analyzed for determining the shape of a substance, as long as the substance could diffract. Not just any substance can diffract light. Certain types of surfaces are better for diffraction than others (e.g., jagged edges diffract better than smooth edges). In 1924, while at Cambridge University, British physicist Jon Bernal solved the structure of graphite and worked on solving the structure of bronze. One of his more distinct and wide-reaching inventions was an *X-ray spectrogoniometer*. A goniometer is an instrument that allows an object to be rotated and measured at different angles. For X-ray crystallography, goniometers are used for measuring angles between crystal facets as well as rotating the sample.

NMR spectroscopy is a technique that measures the energy that resonates in a specific frequency within a magnetic field. It was first developed in 1938 by Isidor Rabi, who later won a Nobel Prize for this invention. It is used to collect physical, chemical, electrical, and structural specifics about molecules. It can provide information about topology, dynamics, and the three-dimensional structure of a molecule in solution or in a solid state. One of its strengths is that it is selective in regard to types of atoms studied within a molecule.

Molecular modeling is the array of methods that are used to investigate the dynamics, surface properties, and structure of biomolecules. Beginning with the modeling of the double helix of DNA (deoxyribonucleic acid, which carries genetic instructions for the development of all known organisms and many viruses), many other molecules and their properties have been modeled such as protein folding, enzyme catalysis, protein stability (thermostability assays), and producing conformational variants (open, intermediate, or closed states).

Patch clamp and single-channel recordings enabled both electrophysiologists and biophysicists to study multiple channels or single-ion channels and determine the effects of these channels on different stimuli. The stimuli could be anything from protons to chemicals, or ligands that are introduced via a special apparatus that controls the rate of flow over a cell. The electrical potential is then measured in amperes and assessed for length of the channel being in open or steady-state conditions. Using these data, dose-response curves can be generated for a particular drug that is introduced to the receptor on the cell's surface.

The *Ramachandran plot* was created in 1963 by biophysicist G. N. Ramachandran and colleagues in order to visualize backbone dihedral angles against amino acids in a protein structure. Today, the plot is generated using X-ray diffraction data for a sample protein and compares the distribution of data points from the sample to that of a known protein structure (structure validation). This task is more easily accomplished if there is a known structure within a family of proteins.

In the late 20th century, *condensed phase single-molecule fluorescence spectroscopy* was devised by W. E. Moerner and Lothar Kador. It uses the fluorescence of a molecule to derive measurements about its structure, position, and its environment. Its strength is that it allows the investigator to obtain information without losing important values due to ensemble averaging. When performing recording techniques, such as patch clamp electrophysiology, the signal that is obtained from the recording of many molecules at the same time is considered the average property of the molecule's dynamics. By focusing on one molecule, it is possible to obtain two-state trajectories.

Nerve Impulse

Following the animal experiments conducted by late-19th-century biophysicists, a prominent English biophysicist, Archibald Hill, performed significant work in the area of nerve conduction and enzyme kinetics. He is known for expanding Michaelis–Menten kinetics and authoring the Hill coefficient; the latter characterizes the binding cooperativity for enzymes and molecules. His work in heat generation by contracting muscles was conducted in 1920. He focused on studying the effects of electrical stimulation to muscle, the mechanical efficiency of muscle, muscle recovery through energy processes, interaction between oxygen and hemoglobin; and produced drug response curves from muscle. He was able to show that electrical activity was most robust in the presence of oxygen, and that during recovery the amount of oxygen required over resting was equivalent to 20% of the lactic acid produced during exercise. This amount of oxygen was needed to resynthesize lingering lactic acid into glycogen.

After World War II, the development of electronics vastly expanded, including radar technologies and the use of nuclear reactors in peacetime to produce large quantities of radioactive isotopes. Alan Hodgkin and Andrew Huxley utilized these advances and data from earlier studies in animal physiology (e.g., Julius Bernstein's membrane theory of nervous conduction), to produce a mathematical model that explains how action potentials are initiated and propagated in neurons (1952). They received the Nobel Prize in Physiology or Medicine in 1963 for such a groundbreaking and important work. Another post-World War II invention, the *electron microscope*, allowed the description of a muscular contraction at the structural level. Many advances in X-ray technology and electron microscopy have made it possible to observe and characterize many molecules that are involved in muscular contraction.

Sensory communication is another area that biophysics has impacted by technological advances that allow scientists to be able to measure stimulus, namely computerized equipment that is far more sensitive than earlier instruments. In general, quantitative analyses of sensory responses are difficult due to the multiple cell types involved. Before the advent of the computer in the 1970s,

Georg von Bekesy, a Hungarian physician, performed experiments with cadavers on the human ear. He dissected the inner ear of the cadaver while leaving the cochlea partially intact. Then he used strobe photography and silver flakes as a marker (which the black-and-white film easily detected) and was able to witness the surface wave of the basilar membrane when stimulated with sound vibrations.

Molecular Biophysics

American scientist Alexander Hollaender founded the study of radiation biology and produced early evidence of nucleic acid as the genetic material. In 1939, Hollaender published data showing that the absorption spectra of nucleic acids of a mutated ringworm fungus was the same as isolated nucleic acids, and therefore concluded that nucleic acids must be the building blocks of genes. However, his discovery was not immediately accepted until other experiments corroborated his hypothesis. It is with this information that a spectrophotometer can be used at a certain wavelength (260 nm) to detect and quantify the amount of DNA in a sample. This type of technology is used routinely in biological laboratories today.

Robert Brainerd Corey, American biochemist, is mostly known for his role in the discovery of the α -helix and the β -pleated sheet with Linus Pauling. In 1950, Pauling and Corey discovered that secondary conformations of peptides were arranged in such a way that their hydrogen-bonding capacity of the backbone NH (amine) and CO (carboxyl) groups was facilitated. First described was the α -helix, a rod-like shape that was composed of a tightly coiled backbone in the inner part with the side-chains extended outward forming a helical pattern. The coil is stabilized by the NH and CO groups of the primary structure. All of the main-chain NH and CO groups are hydrogen bonded. The angles were described in terms of angstroms (distance from one atom to another) and pitch (screw-shaped pattern). Later, the duo discovered another facet of secondary structure, the β -pleated sheet. This pattern differs markedly from the α -helix, taking an accordion-like appearance. In the beta sheets, beta strands, which are polypeptide chains of about three to 10

amino acids long, connect laterally by two to three hydrogen bonds in the backbone.

In 1953, biophysicist Francis Crick and biochemist James Watson determined the structure of DNA, one of the most important discoveries of modern science. Advancements in the use of X-ray diffraction for the determination of crystal structure offered the opportunity to investigate the molecular structure of DNA and RNA. Together the two utilized Chargaff's rule (which stated that for every adenine (A) there is a thymine (T) and for every cytosine (C) there is a guanine (G)) in conjunction with the X-ray crystallization data from Rosalind Franklin and Maurice Wilkins, to determine that the DNA structure was a double helix. There have been some minor changes in the model since it was published, but four of the major features still hold true: (1) DNA is a double-stranded helix, (2) most DNA helices are right-handed, (3) the double helix is antiparallel, and (4) the DNA base pairs are connected through hydrogen bonding, as well as the outer edges of the exposed nitrogen-containing bases. The methods used by Watson and Crick to determine the DNA structure are now used for model building in biophysics.

Donald L. Caspar, an American scientist, made significant contributions in structural biology, X-ray, neutron and electron diffraction, and protein elasticity. In 1962, Caspar and Aaron Klug introduced the concept of quasi-equivalence to account for the arrangement of proteins on the surface of icosahedron virus particles. Caspar-Klug theory has played an important part in shaping the subsequent study of viruses and other macromolecular assemblies. The original concept was based on electron microscopic studies and has now been refined to take into account the atomic resolution structure of viruses, and other details of protein-protein interactions that crystallography has elucidated. Quasi-equivalence continues to be an important component of the philosophical basis for how we think about macromolecular assemblies.

In the 1980s, Adriaan Bax, a molecular biophysicist, developed the field of biomolecular NMR spectroscopy, and he is responsible for many of the standard methods in the field. He pioneered the development of triple resonance

experiments and technology for resonance assignment of isotopically enriched proteins. He was also heavily involved in the development of using residual dipolar couplings (dipole-dipoles) and chemical shifts for determining RNA and protein structures. Recent work focuses on the roles of protein in membranes. He is one of the most cited and important scientists in the field of biomolecular NMR research. Direct dipole-dipole coupling is useful for molecular structural studies. The estimation of the coupling allows for a direct spectroscopic translation of the distance between nuclei and the geometrical form of a molecule. However, in isotropic solution, dipole-dipole couplings are absent as a result of rotational diffusion.

Gary Ackers, Emeritus Professor of Biochemistry and Molecular Biophysics at Washington University, focused on the thermodynamic linkage analysis of biological macromolecules, addressing the molecular mechanism of cooperative O₂ binding in human hemoglobin, beginning in the early 1970s. Briefly, cooperative binding occurs when a substrate binds to an enzymatic subunit (often protein molecules are a series of homogenous or heterogenous subunits that connect on the cell membrane to form one unit) and then activates the remaining subunits. Cooperativity can be positive, negative, or absent. Positive cooperativity results when oxygen binds to hemoglobin. The ratio of oxygen to ferrous iron in a hemoglobin molecule is one-to-one. Deoxy-hemoglobin has a low affinity for oxygen, but when one molecule of deoxy-hemoglobin binds to a heme site, the molecule's overall affinity for oxygen increases. Once this occurs, more oxygen molecules can bind to the hemoglobin molecule. Conversely, in negative cooperativity, when a ligand binds to a protein, the affinity for the ligand will decrease. The affinity for a ligand can be measured by plotting the amount of binding over the course of time.

In standard gel electrophoresis, which separates DNA molecules, there was an issue in that large DNA molecules (larger than 20 kb) would migrate together and it was difficult to obtain adequate resolution or detect the correct size DNA bands. Charles Cantor, an American molecular geneticist, along with David Schwartz, developed pulse-field gel electrophoresis for very large DNA molecules. This technique works by using an alternating

voltage gradient to resolve larger DNA molecules. Opposed to standard gel electrophoresis that runs unidirectional, the pulse-field gel's voltage is pulsed in three directions. The process takes longer than standard gel electrophoresis. Other contributions by Cantor and his lab at Boston University include methods for separating DNA molecules, increasing sensitivity for detection of proteins, and enabling the study of complex protein and nucleic acid relationships.

Pamela J. Bjorkman, an American biochemist, focuses her study on the three-dimensional structures of proteins related to Class I MHC, or major histocompatibility complex, which are proteins of the immune system. Bjorkman and her students study protein interactions involved with immune recognition, using techniques such as X-ray crystallography and confocal or electron microscopy.

Biological Membranes

Membrane biophysics is study of biological membrane structure and function by mathematical, physical, computational, and biophysical methods. In tandem, the data from these methods can be used to generate phase diagrams of various membrane types, describing their thermodynamic characteristics. In contrast to membrane biology, membrane biophysics provides quantitative information and accurate modeling of a variety of membrane topological presentations, such as protein-lipid coupling, lipid and cholesterol flip-flop rates, and lipid raft formations. Also, elasticity of the biological membrane and its intracellular effects can be studied.

One of the methods for visualizing a membrane surface is by attaching a radioactive compound that can either be introduced by insertion by recombinant techniques or added to a solution to bind to a specific receptor. In the early 20th century, the Danish physiologist August Krogh set the foundation for using radioisotopes in molecular biology. His understudy, Hans Ussing, described the pathway that allows the transport of ions across membranes to be explained. He is responsible for the discovery of the active transport of water and ions across a cell membrane. However, the complete molecular mechanism for how this is achieved is still poorly understood.

Despite the great progress in understanding mechanisms of how molecules are assembled into larger biomolecules, there remains a deficit in understanding of exactly how molecules are assembled into biological membranes. The way in which a membrane is constructed is extremely complex and requires better technology to gain useful insight.

Structural Biology

Structural biology is a branch of molecular biology that combines the disciplines biology, biochemistry, and biophysics. The main objective is to obtain a molecular structure of macromolecules, such as nucleic acids and proteins. In addition to structural information, functional data are elucidated through experiments such as electrophysiology. The proteins are usually evaluated in their most complex structures (tertiary and quaternary) and native states because proteins are functional in these states. When a protein is not in these states, inactivation occurs. A great deal of structural biology is centered on determining the structure of receptor proteins that function within a specific disease pathway or has an important physiological role. Often these proteins are crystallized in a specific condition and ligands (or specific compounds that bind to the receptors) are added that will allow the structural biologist to observe the changes in conformation that take place while the ligand is bound. If available, the ligand is a known drug that has a high affinity to the receptor that will promote a physiological response or block it.

Between 1982 and 1985, the German biochemists Johann Deisenhofer, Hartmut Michel, and Robert Huber determined the first crystal structure of an integral membrane protein, the photosynthetic reaction center. It is a membrane-bound complex of proteins and cofactors that play an important role in photosynthesis. They achieved this through X-ray crystallography to determine the exact arrangement of more than 10,000 atoms that make up the protein complex. Not only did their research increased the general understanding of the mechanisms of photosynthesis, but seeing a structure solved for an integral protein was important to help other biophysicists piece together new protocols for their own projects.

Helen M. Berman, an American biophysicist, is the director of the RCB Protein Data Bank. Her biophysics work includes structural analyses of protein–nucleic acid complexes and the role of water in molecular interactions. The Protein Data Bank is an online repository for three-dimensional structural data of large biomolecules (proteins and nucleic acid). The database houses hundreds of protein structures. This online repository is the premier source of data for structural biologists. Most of the data commonly reported with the structure is X-ray diffraction or NMR data. Scientists from around the world are able to submit their structures and are able to download structural data for comparison with their own models. This enables scientists to adjust their models or raise questions about existing structural models.

Steven M. Block (Stanford University) pioneered the use of optical tweezers, a technique developed by Arthur Ashkin, to study biological enzymes and polymers at the single-molecule level. Optical tweezers are instruments that use a highly focused laser beam that attracts or repels according to the refractive index mismatch. This technique has been highly utilized in modern biophysics. The ability to control single molecules with nanometer precision and piconewton accuracy has allowed scientists to directly test the effects of DNA binding on DNA properties and to test the interactions involved in these systems. Understanding the complex mechanisms involved in biological processes is very important and has the potential to improve therapeutics and other molecular biology methodologies.

Related to optical tweezers and single-molecule fluorescence, visualization techniques such as scanning-force microscopy have allowed biophysicists to study the structure and function of nucleoproteins. Carlos Jose Bustamante is an American molecular biologist who pioneered use of optical tweezers and scanning-force microscopy. His laboratory focuses on developing methods of using single molecules to characterize the elasticity of DNA, to induce unfolding of individual protein molecules, and to probe the mechanical behavior of molecular motors.

Axel T. Brünger is known for developing a computer program used for solving structures based on X-ray diffraction data or solution NMR

data. The program is widely used because of its various features. The program is a major extension of the program his team developed called *PLOR*, which came out in 1987. The program utilized a method called *simulated annealing* to refine X-ray crystal structures. It was a manually intense process. When Brünger introduced simulated annealing crystallographic refinement in his 1987 paper, the time to refine crystal structures was significantly reduced and it had a tremendous impact in the crystallographic community.

In some families of proteins, such as the G-protein-coupled receptors (GPCRs), the majority of receptors' structures have been solved. These receptors sense molecules outside the cell and activate signal transduction pathways, which lead to cellular responses. They are difficult to crystallize because there are seven transmembrane receptors that pass through the cell membrane. There are six classes of GPCRs that are homologous and functionally similar. Much information has been obtained regarding their roles in signaling and ligand-binding.

Computational Biology

Predicting the three-dimensional structure of a protein from its amino acid sequence can be a daunting task, especially when there are no homologous structures to compare a new protein structure with. Structure prediction is diametrically opposite from protein design. Protein prediction is one of the most important goals of structural biology. By knowing the structure, it is possible to design drugs that will have greater specificity with the receptor target, thereby causing less side effects overall. Also, enzymes can be created that have biological or industrial uses. In order to help remove some of the obstacles in protein prediction, a number of bioinformatic programs have been created that allows the user to do more tasks at a time.

Some example of molecular design software are Pymol, AMBER, Ascalaph Designer, BOSS, DOCK, Firefly, Maestro, SCIGRESS, SPARTAN, and TINKER. Pymol is an open-source molecular visualization program that was created by Warren Lyford DeLano. It is most commonly used to produce high-quality three-dimensional images of an array of biomolecules. It is one of the only

open-source visualization programs available to structural biologists. It is programmed in the Python programming language.

David Baker is the principal investigator of the 60+ member Baker laboratory. He and his team developed the Rosetta algorithm for *ab initio* protein structure prediction, which is associated with the computing project called *Rosetta@home* and *Foldit*. Members of the Baker group also focus on protein design; they are the first group to have designed a protein, named Top7, with a unique fold. Baker received the 2008 Sackler International Prize in Biophysics for his work in protein folding.

Newer Applications of Biophysics

Astrobiophysics

Astrobiophysics is another emerging application of biophysics that utilizes data obtained from many diverse fields of study, such as biochemistry, paleontology, atmospheric science, and astronomy. Dr. Brian Thomas and his research team at Washburn University (Kansas) apply these principles to study how the Earth is affected by radiation from space. There are several projects underway, most notably the study of the impact on Earth of a nearby supernova. The project uses computational modeling to observe the combined effects of atmospheric ionizations and radiation impact on land and sea life. The presence of the isotope iron-60 in sediment cores bears witness to the fact that a near-Earth supernova occurred within the last 2 million years. The discharge from the event was close enough to change the atmospheric UV levels and possibly modified the Earth's climate. Their work continues in hopes of gaining insight into what sort of impact could a future near-Earth supernova pose to life on the planet.

Medical Biophysics

A fairly new (2013) application of the biophysics discipline is in the field of medicine. There are many classic biophysicists who use methods for determining structures of molecules, membrane surfaces, and proteins in hopes of treating diseases

or better understanding a particular disease pathway. Medical biophysicists employ the findings of these studies in the treatment of patients, in addition to using scanning equipment (e.g., NMR). Individuals who are interested in this new area of study have only a few institutions (Canada and Europe) that offer coursework explicitly aimed at medical biophysics. In the United States, graduate-level programs usually offer PhDs or dual degrees (MD/PhD) in physiology and biophysics, but the interest is growing.

It is easy to see from the history of various disciplines that a consensus on the definition of what biophysics is difficult to achieve. At best, biophysics supplies molecular biology with highly precise quantitative measurements in biophysical processes that could otherwise not be determined. Although its early emphasis was on physiological problems such as nerve conduction and sensory communication (the ear), there was a clear movement toward molecular biology after Watson and Crick used X-ray diffraction data to determine the structure of DNA. It was understood that the future of science leaned heavily toward elucidating the mysteries of biology. It is now the case, however, that all molecular techniques are used in every biomolecular area of study.

The current emphasis for biophysics is in the area of structural biology. The last 20 years of structural biology has been replete with determining novel structures, often the only member of a protein family, and have facilitated the determination of structures in other members of a protein family. The other half of structural biology, the development of novel proteins and enzymes, has implications for cancer treatments or genetic therapies. These therapies and diagnostic methodologies have paved the way for newer fields applying biophysics such as medical biophysics and even astrobiophysics (ecological impacts). These applications seem to be where the future lies for biophysics and likely will be so in the next decade and beyond.

Mandy M. McBroom

See also Biophysics, 19th Century; Biophysics, Contemporary; Chemistry; Health Care Science; Physics, 20th Century; Physics, Thermodynamics

Further Readings

- Ashrafuzzaman, M., & Tuszynski, J. A. (2012). *Membrane biophysics*. Springer-Verlag.
- Brunger, A. T. (1992). Free R value: A novel statistical quantity for assessing the accuracy of crystal structures. *Nature*, 6359, 472–475. <https://doi.org/10.1038/355472a0>
- Caspar, D. L., & Klug, A. (1962). Physical principles in the construction of regular viruses. *Cold Spring Harbor Symposium in Quantitative Biology*, 27, 1–24. <https://doi.org/10.1101/sqb.1962.027.001.005>
- D'A Heck, H. (1971). Statistical theory of cooperative binding to proteins. Hill equation and the binding potential. *Journal of the American Chemical Society*, 93(1), 1A-288.
- Diesenhofer, O., Epp, K., Miki, R., Huber, R., & Michel, H. (1985). Structure of the protein subunits in the photosynthetic reaction center of *Rhodospseudomonas viridis* at 3 Å resolution. *Nature*, 318(6047), 618–624. <https://doi.org/10.1038/318618a0>
- Gilman, A. G. (1987). G proteins: Transducers of receptor-generated signals. *Annual Review of Biochemistry*, 56, 615–649. <https://doi.org/10.1146/annurev.bi.56.070187.003151>
- Glaser, R. (2012). *Biophysics: An introduction* (2nd ed.). Springer.
- Katz, B. (1986). *Archibald Vivian Hill. Dictionary of national biography*. Oxford University Press.
- Liu, Y., Zhang, X., Tan, Y. L., Bhabha, G., Ekiert, D. C., Kipnis, Y., Bjelic, S., Baker, D., & Kelly, J. W. De novo-designed enzymes as small-molecule-regulated fluorescence imaging tags and fluorescent reporters. *Journal of the American Chemical Society*, 136(38), 13102–13105. <https://dx.doi.org/10.1021%2Fja5056356>
- Rasmussen S. G., Choi H. J., Fung J. J., Pardon, E., Casarosa P., Chae, P. S., Devree B. T., Rosenbaum, D. M., Thian, F. S., Kobilka, T. S., Schnapp, A., Konetzki I., Sunahara, R. K., Gellman, S. H., Pautsch, A., Steyaert, J., Weis, W. I., & Kobilka, B. K. (2011). Structure of a nanobody-stabilized active state of the $\beta(2)$ adrenoceptor. *Nature*, 469(7329), 175–180. <https://dx.doi.org/10.1038/nature09648>
- Watson, J. D., & Crick F. H. (1953, April). Molecular structure of nucleic acids: A structure for deoxyribose nucleic acid. *Nature*, 4356. <https://doi.org/10.1038/248765a0>
- Wu, H., Canfield, A., Adihakari, J., & Huo, S. (2010). Quantum mechanical studies on model alpha-pleated sheets. *Journal of Computational Chemistry*, 30(3), 1216–1223. <https://dx.doi.org/10.1002/jcc.21408>
- Wu, H., Wacker, D., Mileni, M., Katritch, V., Han, G. W., Vardy, E., Liu, W., Thompson, A. A., Huang, X. P., Carroll, F. I., Mascarella, S. W., Westkaemper, R. B., Moiser, P. D., Roth, B. L., Cherezov, V., & Stevens, R. C. (2012). Structure of the human k-opioid receptor in complex with JDTic. *Nature*, 485(7398), 327–332. <https://doi.org/10.1038/nature10939>

BIOPHYSICS, CONTEMPORARY

Biophysics is the application of physical concepts and methods toward understanding the living world. As in other branches of physics, experimental biophysical investigation is tightly coupled to the construction of theoretical models that form a mathematical narrative for the phenomena under study, and which help guide and interpret experiments. Whereas 20th-century biophysics largely followed a reductionistic approach, focused on understanding isolated biomolecules and biomolecular processes, contemporary biophysics frequently seeks to understand whole cells, organisms, and communities as self-organized systems, whose properties reflect the combined influence of selective evolutionary pressure and physical constraints.

Theory in Biophysics

Physics aims to form a narrative for all natural phenomena. This narrative takes the form of theoretical models, which are used as everyday tools in physicists' hands to formulate hypotheses and to guide and interpret experiments. Beyond the critical requirement of providing predictive power toward future experiments, a theoretical physical model must exhibit several other characteristics to be considered satisfactory: (a) It must be *quantitative*, expressed in mathematical equations rather than in words, pictures, and so on, which are acceptable formulations for theories in some other disciplines. (b) It must be *simple*. Following Occam's razor, a model is considered better if it has fewer equations and parameters. (c) At the same time, a physical model should also be *universal*, explaining multiple diverse phenomena.

An example of how these requirements can be simultaneously fulfilled is Isaac Newton's law of gravity, whereby the same mathematical equation can be used to calculate the trajectory of an apple falling from a tree, a rocket flying to Mars, or planets orbiting the sun. And in the centuries since Newton's time, theoretical models guided by these same criteria have been developed to describe inanimate matter across scales, from the subatomic to the cosmological. *Biophysics* seeks to extend these efforts to living matter, such that we can understand and predict the behavior of living systems using a quantitative, simple, and universal narrative. It must be noted at the outset that, compared to theories of nonliving systems, the biophysical effort is in an infinitely more primitive state—akin, perhaps, to where the rest of physics was in the 17th century, with (as Newton said) the great ocean of truth lying all undiscovered before us.

In the 20th century, biophysical investigation largely followed a reductionist approach, applying known physical theories to understand the properties of biomolecules, such as DNA, proteins, and lipids, and biomolecular processes, such as the folding of proteins into their functional three-dimensional conformations. In the early 21st century, these theoretical interrogations continue to gain power owing to rapid progress in enabling experimental techniques, such as cryogenic electron microscopy and fluorescence microscopy, and computational methods, such as molecular dynamics simulations and machine learning tools. Altogether, these advancements enable molecular biophysics to move from the study of isolated molecules in an aqueous solution to examining more elaborate scenarios, where multiple molecular species interact in a chemically complex environment, thus beginning to emulate what takes place in the living cell rather than in a laboratory test tube.

The Physics of Living Systems

A complementary approach to biophysics, which was always present in the field but gained prominence in the 21st century, focuses not on the isolated constituents that make up living matter, but rather on how these constituents come together to form *living systems*—cells, organisms, and communities. This effort is inspired by, and

aims to emulate, the success of statistical physics to predict how macroscopic properties of inanimate matter emerge from the interactions between its microscopic constituents. This theoretical success was originally limited to simple materials that are homogenous and orderly, such as crystals, but gradually grew to include more complex systems that are heterogenous and disordered, such as glass, or out of equilibrium (i.e., undergoing change) such as weather phenomena. Extrapolating from this success, living systems are viewed as the ultimate emergent phenomena, and understanding them as the natural next step in the study of complex systems.

Theories in this area of biophysics (sometimes called *physics of living systems*) thus seek to explain how biomolecules self-assemble into machines, complexes, and networks that underlie the function of the living cell; how individual cells likewise organize to form a brain, an immune system, and the other multicellular structures that make up a higher organism; and how organisms again self-organize into communities and ecosystems. Moreover, beyond the understanding of living systems as a subject matter, developing such theories will lead (it is hoped) to the discovery of new physical principles—*laws*—governing the process of self-organization. Thus, this branch of contemporary biophysics goes beyond the traditional strand by anticipating new physics from biology, rather than only new biology through physics.

But applying the statistical-physics paradigm toward understanding biology faces considerable obstacles, reflecting the unique properties of living systems as compared with their inanimate counterparts. First of those is the absence of *scale separation*. In traditional physics, different spatio-temporal scales can often be considered separately. For example, when describing the flow of liquid at the macroscopic scale, one can ignore the atomistic nature of material at the microscopic level. This conscious ignorance is enabled by the theoretical practice of coarse graining, where a multitude of details on a lower scale (e.g., the complex interaction between nearby molecules of the liquid) is mapped to much fewer observables on the higher scale (the resulting viscosity of the liquid). Matters are very different in biological systems, where critical information is

typically present at all spatial and temporal scales, from the conformational changes of individual biomolecules at the nanometer and nanoseconds range to the resulting changes in characteristics (phenotype) of a multicellular organism, taking place over meters and hours. Consequently, there is no standard procedure for coarse graining over the molecular details of a living system when aiming to describe it at the cellular, or higher, level. This poses a significant obstacle on the way to constructing a simple theoretical narrative in the physical tradition.

To make matters worse, the detailed information required for constructing a biophysical model is simply unavailable for the vast majority of molecular processes in the cell. Even in the best studied organisms, which serve as *model systems* for the rest of the living world, what typically exists is a partial list of molecular players involved in a given process (e.g., the transcription of RNA molecules from their DNA template), but none of the biophysical parameters—binding affinities, reaction rates, and so on—that characterize the underlying molecular interactions. Even for the handful of select processes on which researchers have focused over decades—for example, how the activity of a specific sugar utilization gene in *Escherichia coli* bacteria is regulated—what is mostly available are measurements made on isolated molecules in a laboratory test tube. The relevance of these measurements to the complex intracellular environment, where physical conditions are very different, and where numerous additional players are present, is uncertain. The bottom line is that, even if building a detailed molecular model for everything in the cell were computationally feasible, the information required for constructing such models is still missing.

Minimal Mechanistic Models and Occam's Rug

To circumnavigate these obstacles, simple theoretical models for biological phenomena are often constructed by trying to identify a minimal set of features that are essential for the observed phenomenon—and are therefore explicitly included in the mathematical representation—while ignoring many other details. The choice of which details to

ignore reflects a conscious decision by the biophysicist, but also, invariably, the fact that (as noted above) those details are largely unknown. In addition, identifying which features to include in the model is often done via trial and error by testing what choices improve the agreement between theoretical predictions and experimental observations. Owing to this targeted process of model construction, such models are more properly considered phenomenological (top-down) rather than truly reductionistic (*bottom-up*). Another hallmark of these simplified models is referred to as *Occam's rug*: In contrast to the physics of nonliving systems, where a simple theory is thought to bear the hallmark of truth (Occam's razor), in the case of biophysics, a model's simplicity instead implies that most of the details have been swept under the rug.

The minimal modeling approach has proven powerful when applied to biological phenomena that are relatively simple and well characterized, with the canonical examples coming from *E. coli*, the best studied of all model organisms. Mechanistic biophysical theories have been formulated to capture the way a bacterial cell senses changes in its chemical environment and modulates its behavior accordingly, by swimming toward nutrients and by producing the enzymes required for consuming these nutrients. Following the earlier success in bacteria, the first decades of the 21st century have seen an explosion of minimal mechanistic models directed at systems of higher biological complexity, including processes taking place in eukaryotic (nucleated) cells, during embryonic development of multicellular organisms, and during the onset of human disease, such as the emergence of cancer.

Unfortunately, evaluating the success of biophysical models is less straightforward than it is for models of inanimate matter. This is because, as noted above, the construction of biophysical models involves cherry-picking what features to include, and, consequently, the resulting models are rarely unique mathematically or biologically. In other words, the fact that a given theory recreates existing data does not necessarily mean that it correctly captured the essential biological features. This problem is made more severe by the fact that matching theory and experiment almost invariably involves the process of fitting unknown

mathematical parameters. These parameters represent biophysical properties of the system, such as molecular concentrations and diffusion and reaction rates, but since their values are unknown, they become ad hoc correction factors that artificially increase the model's chances of success. This weakness of mechanistic biophysical models often reveals itself in their limited predictive power outside the narrow premise for which they were calibrated, thus failing the premise of universality desired in physical theories.

Modeling Cellular Individuality

Both the success and limitations of minimal biophysical models are demonstrated by the question of single-cell behavior. It has long been observed that individual living cells, even if they are genetically identical to each other and placed in the same environment, nevertheless may exhibit large differences in the expression of their genes and, consequently, their phenotype. This cellular *individuality* has important consequences across biology, from the emergence of resistance to antibiotics among bacteria to the choice of cellular identity during early embryonic development. The individuality of cells places an additional challenge for biophysical models: beyond capturing the average behavior of a cell population, also predicting the cell-to-cell differences in this behavior.

Attuned, perhaps, to the contemporaneous zeitgeist emphasizing heterogeneity over uniformity, biophysical theory turned its attention to cellular individuality during the late 20th century. To do so, it called on the concept of *noise*, as used in the study of electronic communication, where it indicates random, unwanted fluctuations that accompany the transmitted signal and may obscure it. In the cellular context, noise—the deviation of single-cell properties from the expected average behavior—arises from the inherent randomness of biomolecular events such as diffusion and binding, which in turn results in cell-to-cell differences in the production of regulatory molecules and, consequently, in phenotype. The physical theory of stochastic processes provides the tools for calculating the expected degree of cellular heterogeneity, in the form of the statistical distributions of any quantity, for example, the numbers of proteins

produced from a gene. This theoretical approach has proven successful in reproducing experimental measurements of gene expression in individual bacterial cells, and it has since been expanded to describe other facets of cellular individuality.

But utilizing the concept of *noise* to describe cellular heterogeneity also comes with a strong caveat. A key implication of the stochastic picture is that single-cell behavior is inherently unpredictable, and that a living cell's every choice is subject to significant randomness. However, by implying that single-cell observables are not merely unknown but also *unknowable* (in that they are unpredictable), the noise idea legitimizes our ignorance of these cellular properties, sweeping them too under Occam's rug. A stochastic description of cell behavior may thus create a facade of understanding, while in fact impeding the efforts to reveal deterministic drivers of that behavior. Conversely, forgoing the assumption of randomness and identifying those *hidden variables* (a term borrowed from quantum mechanics) will allow biophysical theories to regain predictive power at the single-cell level. The search for hidden variables that drive cellular individuality is an active area of study.

Beyond Mechanistic Models: Evolved Optimality Under Physical Constraints

A complementary approach to developing a theoretical description of living systems posits that these systems' overwhelming complexity, and our considerable ignorance of the underlying molecular details, should argue against the construction of mechanistic models. Instead, biophysical theories should leverage the one universal law of biology already known to us, namely, evolution. Specifically, the premise of theoretical analysis should be that the observed properties of any biological system reflect the outcome of selective pressure to optimize its function, as the system evolves under a set of physical constraints. This investigative approach, typically referred to as *systems biology*—although that term carries different meanings to different people—is highly interdisciplinary. In addition to physics, it leverages theoretical tools from fields as remote as information theory (originally used to describe the

communication of digital information), control theory (used to describe human-made devices), economics, and other areas where the question of optimal performance in the presence of external constraints is often of interest.

Systems biology has had multiple successes in identifying and explaining the properties of living systems. One example is the discovery of common topological features in biological networks. These network *motifs*—typical patterns of connectivity between network nodes—optimize the processing of information by the network. Strikingly, the same motifs are found across systems as disparate as the genetic circuits within a cell and the neuronal networks in the brain, as well as the (human created) World Wide Web, all of which share the requirement to transmit information efficiently to function properly. Other types of motifs endow networks with the property of *robustness*, defined as the ability to function in the presence of uncontrolled fluctuations of both the environment (e.g., changes in temperature) and of the system's internal components (e.g., varying numbers of cellular proteins). Applying principles from control theory led to the identification of robustness in the biochemical networks that underlie bacteria's ability to sense and swim toward nutrients, and in genetic circuits that drive the early stages of embryonic development. In another example of using a systems biology approach, researchers utilized the economic principle of *Pareto optimality*, which describes how trade-offs between different tasks can be best reconciled, to explain several quantitative relations observed across biology, from the shapes of Darwin's beak of the finch, and the sizes of bat wings, to the sets of genes bacteria express under changing growth conditions.

As exemplified by the cases above, the theoretical models of systems biology often constitute a satisfactory physical narrative, in being relatively simple but at the same time universal—that is, applicable across diverse biological instances. But this attractiveness does come with a price tag. For one, while a self-consistent explanation for the observed phenomenon is offered, the underlying mechanistic details (the *how*), which invariably differ between biological systems, often remain unknown. Another aspect sometimes missing from these theories is novel predictions regarding future

experiments; for example, the expected response to specified perturbations. Such predictions are required as means to falsify the proposed theory, which, in their absence, risks becoming more of a *Just So* story.

Concluding Remarks

In evaluating contemporary biophysical theory, it must be noted that, even for the best characterized living systems, we are still very far from achieving a comprehensive *molecules to organisms* narrative that fulfills the simultaneous requirements of quantification, simplicity, and universality. An honest assessment of the field must question whether such models are even an achievable goal. Put differently, beyond important but limited *niches*, such as illuminating the function of biological molecules, what roles can physical theories play in understanding the living world? Of course, one may simply argue that physicists' interest in pursuing theoretical interrogation of living systems is sufficient justification for the endeavor. After all, we must remember that intellectual curiosity has always been the driver of discovery. Moreover, the biophysical endeavor has significant pedagogical value, as it involves the training of interdisciplinary scientists capable of applying rigorous quantitative analysis to biological and medical problems. On the other hand, if one considers that the role of theory is, above all, to provide predictive power—here, forecasting the future behavior of a living system under previously untested conditions—then, in the 21st century, it may be argued that a more promising avenue to achieving that power is offered by machine learning, whereby a computerized algorithm learns from past information to predict future data, without going through the *middleperson* of a human-curated model. But a more balanced view suggests that model-free prediction can provide benchmarking for the performance of human-created models, benchmarking which physicists never had before. Furthermore, machine learning methods can aid in the construction of theoretical models, by helping to delineate critical from ignorable features, and to identify model rules that predict experimental observations. Thus, rather than harbinger of a future of *prediction*

without understanding, machine learning methods promise to become an invaluable tool in the theoretical arsenal of biophysicists.

Ido Golding

See also Biophysics, 20th Century; Cell Theory; Kinetic (Molecular) Theory; Modeling; Natural Selection; Physics, Contemporary; Systems Science

Author's Note: Ido Golding is supported by grant R35 GM140709 from the National Institutes of Health.

Further Readings

- Alon, U. (2019). *An introduction to systems biology: Design principles of biological circuits* (2nd ed.). CRC Press.
- Berg, H. C. (2004). *E. coli in motion*. Springer.
- Bialek, W. S. (2012). *Biophysics: searching for principles*. Princeton University Press.
- Bishop, C. M. (2006). *Pattern recognition and machine learning*. Springer.
- Dill, K. A., & Bromberg, S. (2011). *Molecular driving forces: Statistical thermodynamics in biology, chemistry, physics, and nanoscience* (2nd ed.). Garland Science.
- Golding, I. (2011). Decision making in living cells: Lessons from a simple system. *Annual Review of Biophysics*, 40, 63–80. <https://doi.org/10.1146/annurev-biophys-042910-155227>
- Ingalls, B. P. (2013). *Mathematical modeling in systems biology: An introduction*. MIT Press.
- Nelson, P. C., Bromberg, S., Hermundstad, A., & Prentice, J. (2015). *Physical models of living systems*. W.H. Freeman & Company.
- Phillips, R., Kondev, J., & Theriot, J. (2009). *Physical biology of the cell*. Garland Science.
- Swain, P. S. (2016). *Lecture notes on stochastic models in systems biology*. <https://doi.org/10.48550/arXiv.1607.07806>

BIOPSYCHOSOCIAL MODEL

The biopsychosocial model requires that biological, psychological, and causal factors causally interact with each other. Hitherto this has been hard to theorize, but the current life and behavioral

sciences show how within- and cross-domain causation works. The biopsychosocial model is a model for medicine proposed by the physician and psychoanalyst George Engel at Rochester, New York, in a series of papers in the late 1970s and early 1980s. The main contrast is with the biomedical model, which focuses on biological factors in disease, while the proposed extension broadens the scope to include psychosocial as well as biological factors. The biopsychosocial model and its comparisons and contrasts with the biomedical model raises many questions for theory across the range of the relevant sciences: the life sciences, the behavioral sciences, and the social sciences, as well as the environmental sciences.

Medicine is an applied science making extensive use of technology for detection and treatment, of many kinds, ranging from physics-based technologies to psychosocial. The biopsychosocial model can also be used as a frame for health professional–patient relationships, emphasizing the importance of taking account of the person as a whole. Although primarily a model for medicine, the biopsychosocial model arose at the time of revolutionary changes in the life and behavioral sciences. Centuries-old assumptions about physical nature, mind, and causation were being overturned, being replaced by foundational concepts such as information processing and systems science. A consequence of these changes is that what was hitherto a disunity among the sciences has given way to increasing recognition of the need for cross-disciplinary research programs.

Biomedical and Biopsychosocial Models

The Basic Contrast

Biomedicine is a set of basic biological and clinical science research programs with associated technologies for detection and treatments. It has made great advances in the understanding of many diseases and in the treatment of many, most spectacularly for infectious diseases, but many others besides. The core basic science is physiology,

increasingly enhanced by advances in the understanding of basic cellular processes and genetics. For practical purposes, biomedicine hardly needs a general model, but it does have the working, methodological assumption that diseases can be understood only or primarily in terms of biological factors. This assumption can be called the *biomedical model*. In 1977, George Engel proposed a new model for medicine, the *biopsychosocial model*, arguing that the biomedical model was too narrow in its focus on biological factors, and that, rather, medicine needed to take account of the full range of biological, psychological, and social factors.

Multiple Issues Involved: Etiology, Treatment, Boundaries

The issues involved in the biopsychosocial model are large, abstract, and complex. The broad question whether psychological and social factors as well as biological factors are involved in health and disease arises in many contexts: the etiology of disease, course of disease post-onset, complications by further health problems, and treatments. In addition, there is the question of how we draw the boundaries between what is disease and what is within the normal range of variation. In each of these cases the question arises, Are the main factors and considerations biological, or are they rather broader, biopsychosocial? All of these issues are of major importance in the health sciences and health care, and they are all in play in Engel's papers proposing the biopsychosocial model. This gives rise to significant ambiguity as to exactly what is being proposed, which is likely one reason for the model's mixed reception, as discussed in what follows.

Deep Theory Also Involved: The Nature of and Relation between Biological, Psychological, and Social Factors

Further, Engel's papers penetrate into deep theoretical issues, including whether biology is reducible to physics and chemistry; the need to abandon the dualism of mind and body; and the importance of causal interactions between biological,

psychological, and social factors. These theoretical issues will be taken up later in this entry.

Mixed Reception of the Biopsychosocial Model

Widespread Acceptance

It would be reasonable to say that the biopsychosocial model is the dominant headline model of the contemporary practice of health care. It is endorsed as obvious by many health scientists, clinicians, and health educators, and as especially important for quality of clinical care and health care education. It is obvious enough that the patient as a whole has a biology and psychology and social life, and that this is important to bear in mind when treating.

Radical Criticisms of the Biopsychosocial Model

Alongside widespread assent to the biopsychosocial model, however, amid its apparent obviousness, another theme in the reception of the biopsychosocial model is that it is vague and useless. The criticism is that the model is mere hand-waving, that it makes no specific scientific claims and gives no specific clinical guidance but can be used to justify anything or nothing. Thus, we have the intriguing picture as a whole that the biopsychosocial model is widely endorsed, but at the same time criticized for being too general and vague to be of any use.

Maybe the Biomedical Model Is Enough?

In the meantime, biomedicine continues to make substantial advances in the treatment of many diseases and continues to attract a high proportion of health research and clinical funding. At the time of writing, Fall 2020, promising interim results of the first to report Phase 3 trial of a vaccine for COVID-19 have just been released; the latest example of great success for biomedicine. More generally, the two most rapidly expanding basic health care sciences, genetics and neuroscience, both appear to be biological, insofar as they focus on biological factors—issues picked up in

the last section. In the context of the continuing success of biomedicine, it can be readily acknowledged as obvious that the person as a patient has a psychological and social life as well as a biological body, and that this has to be kept in mind when managing an illness, while noting that this can be so even if the core disease process is primarily biological, requiring a biological intervention.

The Need for Evidence

Following the preceding line of thought, the biopsychosocial model only presents as a genuine alternative to, actually as an extension of, the biomedical model, inasmuch as it can be shown that psychosocial factors as well as biological factors are causally relevant to the onset and course of diseases. Such evidence has been gradually accumulating over recent decades as reviewed in the following section.

Evidence Base

New Group Research Methodology

The biopsychosocial model was proposed at a time, the late 1970s and early 1980s, when there still tended to be competing theories about the nature of illness in general—that it was caused by one or other of biological, psychological, or social factors. Research was confined mainly to case studies. It was not until the closing decades of the 20th century that randomized controlled treatment trials came into widespread use. Treatment trials are in a class of controlled research designs and associated statistical analytical modeling capable of estimating the proportion of variance in a health outcome of interest attributable to particular factors. The same approach has been applied in epidemiology to investigate risks for development of disease. These new research methodologies have been used to study empirically and quantifiably the role of biological, social, and psychological factors in the onset and treatment of diseases.

Evidence Base for Psychosocial Factors

Findings in epidemiology over the past few decades have included identifying the so-called social determinants of health. The novel finding

has been that social disadvantages are implicated in the onset of many kinds of health problems, physical and mental. There are also many clinical trials showing the importance of psychological treatments in affecting (a) the course of mental health conditions and (b) the adjustment and mental health complications of long-term physical health conditions.

Evidence Base Is Typically Condition- and Stage-Specific (Not “General”)

As these findings have accumulated, they have lent clear support to the biopsychosocial model. That said, the research evidence is specific rather than general; it distinguishes between conditions and stages of condition. In advanced stages of some physical illnesses, such as Ebola virus disease, or Grade 4 cancers, or advanced cardiovascular disease, the driving processes are typically biological only, and treatments, if there are any, are accordingly biomedical. On the other hand, if one looks for risks of disease, whether they be lack of environmental hygiene, or unhealthy diet, or chronic stress associated with economic disadvantage, the causal factors broaden out into biopsychosocial. This broader picture also applies to adjustment and quality of life in chronic health conditions. There is no biomedical cure for long-term conditions, though biomedical management may be critical, and questions of lifestyle and quality-of-life implicating psychological and psychosocial factors come to the fore.

Practical Implications of the Biopsychosocial Model

Public Health

The new findings in social epidemiology imply that reduction in population prevalence of disease would require policy to address chronic stress, with its associated biological damage, by alleviating socioeconomic disadvantages. Public health in this new context is closely linked to economic policy, and generally to policy to alleviate social exclusion from health-related resources such as clean air, healthy diet, and a living wage. Psychosocial factors are also increasingly emphasized in public health strategies to manage infectious disease epidemics, such as population cooperation

with infection control measures and high-enough vaccine uptake.

Clinical Treatments

Treatment trials for many health conditions have suggested efficacy of a range of psychosocial treatments. Such findings lend support to the biopsychosocial model. However, the biopsychosocial model is general and is no substitute for the detailed findings of treatment trials. We now have systematic reviews of treatment trials that summarize the findings for a particular condition or for a particular stage of a particular condition. These reviews discriminate between the scientific quality of the trials and make recommendations as to management and treatment. These recommendations, or treatment guidelines, always come with the qualification that each patient and presentation is unique, having a unique combination of factors that cannot all be taken into account when inferring an individual management plan from a range of group treatment studies; and that, accordingly, individual clinical judgment in deciding a clinical management plan is required. Treatment plans may be biological, psychological, or social (as in so-called social prescribing), alone or in some combination, typically requiring a multi-skilled, multidisciplinary team. Good clinical care of the patient always involves respecting the whole person.

Theorizing Biological, Psychological, and Social Factors

Dualism and the Reduction

Problem: Material Causes Only

At the theoretical core of the issues considered so far are the questions of what biological, psychological, and social factors are, and how, if at all, they interact. The life and human sciences have developed over the past or so against the background of mind/matter dualism that created deep theory prejudices: that nature consisted of physical matter and physical causes alone, and mind, the realm of cognition, was either noncausal or in some way identical to the material brain and behavior. As to social reality and causes, in this picture they are hardly conceivable. In the 1970s, it was commonly and still reasonably believed that biology was

reducible to physics and chemistry. The stunning successes of biomedicine in identifying the biological causes of many diseases thus coincided with the deep theoretical or philosophical assumption that anyway all causation had to be biological, ultimately a matter of physics and chemistry. Engel proposed the biopsychosocial model in the 1970s against the background of the deeply antithetical assumptions that psychological and social causes were highly problematic ideas, whereas biological causation alone seemed legitimate. The proposal, however, was a signal of what was to come; around the very same time interrelated revolutions in biology and psychology were in the process of undoing the traditional assumptions.

What Is Biology? A New Science of Regulatory Control

A key theory contribution to revolutionizing biology was made by the physicist Ernst Schrödinger, famous for his work in quantum mechanics, in a lecture delivered in Dublin in 1943, in which he proposed that life could be understood as a local system in which the overall direction of the second law of thermodynamics ran, temporarily, in reverse: in biological systems, entropy is decreased; order, energy differences, increased. Schrödinger saw further that this extraordinary feat was accomplished by genetic coding—a new concept that remained to be worked out. Ten years later Watson and Crick published their work on the structure of DNA and the genetic code. From around that time, biology, physiology, including biomedicine, has developed as a balanced combination of two sciences: the biophysics and biochemistry of energy exchanges, for example underlying basic cellular metabolism, but this combined with a new science of regulatory control of those processes based on information exchange or communication. Current biomedicine is in large part a study of regulatory mechanisms regulating metabolic processes and other regulatory mechanisms. This new science has logical formulae for its mathematical representation: regulatory mechanisms are basically stop/start, represented by negation, and dependent on conditions, represented by “if . . . , then. . . .” Information processing science interweaves with computing.

The new framework of causal explanation in biology in terms of information- or communication-based regulatory control is also applicable in the behavioral and social sciences. This enables a more unified science in which there can be causal interactions between these previously disparate domains.

The Psychological as Embodied Agency and the Social as Resource

The new information-processing paradigm also revolutionized psychology at around the same time, in the so-called cognitive revolution in psychology. While early computational models of mind in the 1970s were computerlike, emphasizing only manipulation of symbols, increasingly, as the research program has evolved, the mind is seen as more biological, embodied, in close interaction with the environment. These theories of embodied cognition were first recognized in the tradition of philosophical phenomenology but are now at the cutting-edge of neuroscience. Important for the purpose of unifying the science, and implied by the common use of the information-processing paradigm, there are information-based regulatory pathways between the mind–brain and some biological systems below the neck, linking neuroscience with biomedicine. The mind–brain also regulates behavior in the environment; accordingly, agency as the capacity to act is an overriding function. Regulation is a core feature of social causes, as well as of biological and psychological causes. The environment for us includes the whole range of factors—physical, chemical, biological, and social—and our biological and psychological functioning depends on it. The biopsychosocial model needs to include the environmental sciences.

Causal Interactions

The core consequence of these theory shifts in the basic sciences are that causal interactions are possible between what were previously disparate domains, which now become linked in a unified ontological–causal space. A sign of this in health sciences are multifactorial models of causal

interaction, such as between genes, lifestyle, and environmental exposure, in pathways to many health conditions. This development in the health sciences is part of a broader recognition of the need for problem-solving research programs to be multidisciplinary.

Derek Bolton

See also Health Care Science; Medicine, 20th Century; Medicine, Contemporary; Philosophy of Mind

Further Readings

- Bolton, D., & Gillett, G. (2019). *The biopsychosocial model of health and disease. New philosophical and scientific developments*. Springer Palgrave Open Access. Available at <https://www.palgrave.com/gp/book/9783030118983>
- Bolton, D. (2020). The biopsychosocial model and the new medical humanism. To appear in French translation in *Archives de Philosophie*, Special Issue, J. Ferry-Danini & E. Giroux (Ed.). English version at <https://www.cairn-int.info>
- Engel, G. L. (1977). The need for a new medical model: A challenge for biomedicine. *Science*, 196, 129–136.
- Engel, G. L. (1980). The clinical application of the biopsychosocial model. *American Journal of Psychiatry*, 137, 535–544.
- Frankel, R. M., Quill, T. E., & McDaniel, S. H. (Eds.). (2003). *The biopsychosocial approach: Past, present, future*. Rochester University Press.
- Ghaemi, S. N. (2010). *The rise and fall of the biopsychosocial model: Reconciling art and science in psychiatry*. Johns Hopkins University Press.
- Lehman, B. J., David, D. M., Gruber, J. A. (2017). Rethinking the biopsychosocial model of health: Understanding health as a dynamic system. *Social and Personality Psychology Compass*, 11(8). <https://doi.org/10.1111/spc3.12328>
- Marmot, M. G. (2006). Status syndrome: A challenge to medicine. *Journal of the American Medical Association*, 295(11), 1304–1307. <https://doi.org/10.1001/jama.295.11.1304>
- Schrödinger, E. (1944). *What is life?* Cambridge University Press.
- White P. D. (Ed.) (2005). *Biopsychosocial medicine: An integrated approach to understanding illness*. Oxford University Press.

BIOSTATISTICS

Biostatistics is both a practice and a profession. As a practice, it involves addressing biological or medically related themes using quantitative reasoning. As a profession, biostatistics can be understood as a different type of practice, one that is widely institutionalized, primarily in North America and Europe, with its own societies, associations, journals, commonly taken-for-granted skills and credentials, and applications. After describing this basic terminology, this entry provides context for both the practice and profession of biostatistics as a series of three historical epochs: (1) its *gestation phase*, which largely began in the 19th-century Europe and the United States; (2) its *foundation phase*, whereby in the early 20th century, renewed debate over the nature of evolution led to one of the most historic accomplishments in science: the consensus across a broad array of disciplines in the process of Darwinian natural selection; (3) its *institutionalization phase*, in which the practice and profession of biostatistics became more widely institutionalized, particularly in the United States, along with a larger process of societal medicalization. The entry concludes with a discussion of some of the promises and perils for biostatistics in the 21st century.

Terminology: The Duality of Biostatistics

This section makes two main points. First, that the term *biostatistics* is semantically vacuous: most definitions of biostatistics have either been terse and limited to a one-sentence description, or definitions of biostatistics have been stretched to a dizzying array of domains. To clarify the first point: Biostatistics encompasses a wide variety of statistical applications to biological phenomena. Beyond this statement, further clarification is warranted.

Second, I claim that understanding biostatistics can be clarified further by distinguishing biostatistics as a practice, or *what some people do*

and biostatistics as a profession, or *what some people are*. In other words, biostatistics as a concept can be purposefully widened beyond the tight definitional circularity in which it presently finds itself.

What Is Biostatistics?

The term *biostatistics* is vacuous because, as a concept, it is too broad to be very meaningful. Its sacrificial breadth is not unreasonable to compare it to Jorge Luis Borges's famous short story about a map that was made to such exacting detail that it became as large as the empire it was intended to represent. This *map-territory* problem is a conceptual problem in clarifying the concept of biostatistics: as a construct, it is too encompassing to be very meaningful without being circular semantically. The concatenation of "bio" with "statistics" is most accurately described as a "shotgun wedding at best" (by Morgan, 1986, p. 1105), primarily because it does little to clarify in specific terms or in specific contexts the practice of biostatistics. To define biostatistics as the application of statistics to biological phenomena does not clarify much if the reader knows the word root *bios* is Greek for *life* and that the word *statistics* refers broadly to the classification, collection, analysis, and inference of numerical facts or data. Several exchangeable terms predate the term *biostatistics*. For example, the first recorded use of the term *biostatistics* in the third edition of a dictionary of "medical lexicon," by notable English and American physician Robley Dunglison, refers the reader to the earlier term *medical statistics*, which was used in the volume's 1842 second edition (1865, pp. 138, 654). Even before the concept of biostatistics there were a plethora of other concepts that were used interchangeably with the term *biostatistics*, but which mean different things, including *biostatics*, *biometry*, *biometrics*, *biological statistics*, *medical statistics*, *biostatistical science*, and *biomedical sciences*. More recent terms include *environmental statistics*, *pharmaceutical statistics*, and *public health statistics*.

The Practice/Profession Distinction

To clarify the concept of biostatistics, we need to distinguish between the term as referring to a *practice* and as referring to a *profession*. To preview, biostatistics as a practice emphasizes what is meant when one *does biostatistics*, such as how a question is framed, what topic is addressed, and the procedure or method employed. In contrast, biostatistics as a profession focuses on how people *be biostatisticians*; that is, on how biostatistics has developed as discipline with its own societies, journals, and its own linkages with governmental institutions.

Biostatistics as a practice is relatively ambiguous; however, intuitively, the word itself references both the *statistics* (*statistics* broadly speaking, but more fundamentally a positivist approach toward knowledge construction) and *biology* (*bio*, referring to anything biological). The practice of biostatistics has varied over time and space in its application. About the only unifying methodological approach is its emphasis on quantification. While the practice of biostatistics is difficult to define, greater traction is possible by viewing it as a profession. In this context, professionalization refers to the social processes by which an occupation or a particular status is symbolically and socially constructed.

Beyond these two categories, biostatistics may refer to any number of data analytic techniques and may include any number of applications. The following are only a few topical examples, such as: the study of population prevalence of births and deaths (calculation of so-called *vital statistics*), the study of population health (*epidemiology*), the distributions and changes of genetic diversity of a population (*population genetics*), the sequence of *genomes* or genetic material of an organism or organisms (*bioinformatics*), the study of drug action (*pharmacology*), not to mention applications in agriculture, ecology, animal studies, health economics, and botany, among many other fields. The best way to understand biostatistics is to likely paraphrase famous statistician John Tukey's line: While statisticians "get to play in everyone's backyard," biostatisticians get to play in everyone's "garden," focusing not on the space as a play area, but rather on its flora and fauna.

Another point to be made is that, in large part because of the field's influence on statistics in general, while there are some statistical applications more commonly applied in biostatistics, such as sequence analysis, the majority of biostatistical methods could apply to almost any discipline, although the terminology may differ. For example, while biostatisticians refer to as *hazard* or *survival analysis* as a technique to understand the timing of an outcome (such as "number of years lived before death"), this same technique is sometimes referred to in economics as *duration analysis* and in history and sociology as *event history analysis*.

It is perhaps not surprising that most biostatisticians focus primarily on methods and on data applications. For example, in Wayne Daniel's accessible (and recommended) introduction, *Biostatistics: A Foundation for the Health Sciences*, nearly all of the textbook focuses exclusively on statistical applications, using examples in the health sciences, but with only a brief definition of the term on the third page. Accordingly, the major specialty journals such as *Biometrika*, *Biostatistics*, *Statistics in Medicine*, and *Pharmaceutical Statistics* tend to address new methodological innovations and findings from particular analyses, avoiding the ontological trap of attempting to define the subfield.

The Genealogy of Biostatistics in Three Acts

This section traces the *genealogy* (to borrow Michel Foucault's term) of biostatistics to understand its present incarnation as a historically situated phenomenon. By sticking to historical facts, one can begin to understand biostatistics *not* as I conceitfully introduced it—as a term so vacuous as to embody everything—but rather to understand its *specificity*.

Although biostatisticians are apt to identify pre-19th century, canonical figures, it is important to consider that biostatistics codeveloped with science more broadly and statistics and biology in particular. Biostatistics is best understood as a 19th-century phenomenon, closely aligned with progressivism and with eugenics in the early 20th century, but primarily an established discipline in

the post–World War II period in the United States. As a practice and as a profession, biostatistics has made great contributions to statistics, science, as well as population health. While biostatistics as a practice will undoubtedly continue, there are important questions regarding the future of the profession, in part due the inchoate boundaries of the discipline combined with rapid advancements in technology.

Gestation (1800–1900)

First, and perhaps most importantly, there was the rise of the bureaucratic nation-state, which created an avalanche of printed numbers that required bureaucratic thinkers to both consolidate and manage each state's *informational capital*. By the early 19th century, France, Germany, and the United Kingdom created governmental bodies devoted to collecting statistics on their citizens. A notable example in the United Kingdom was the transition from the church of vital records (e.g., births, deaths, marriages) to the government. In particular, in 1836, the United Kingdom established the General Register Office to manage to supply all vital records and assigned physician William Farr to manage the records. A centralized governmental agency took over the role of the Church of England in large part because of the increasing number of nonparticipants and Catholics in the United Kingdom.

Second, in large part because of the rise of the nation-state, European scientists and philosophers began rejecting the philosophy of causal determinism in favor of a theory of probabilistic thinking that provided an understanding of the world that made it seem less disorderly. Causal determinism was the view that the universe is established mechanistically, with every event occurring in a lawlike fashion in a predetermined causal chain of events. In contrast, probabilistic thinking provided thinkers to develop knowledge that focused on variation among populations and to think of social problems in terms of chance or risk. Bureaucrats armed with an *avalanche of numbers* used such resources to generate a philosophy of chance that guides much of statistical thinking today.

Third, the 19th century underwent political and social movements that promulgated *progressivism*, a broad philosophical idea that asserts that human advancement is made possible through technological economic development in conjunction with the practical application of scientific principles. Progressivism has its roots in 17th-century European Enlightenment ideals emphasizing reason, individualism over tradition, and with such famous French and English thinkers as Francis Bacon, John Locke, Voltaire, David Hume, and Isaac Newton. Although a review of such a grand concept is beyond the scope of this analysis, it is important to underscore that early thinkers were largely (and in many respects) are at least implicitly influenced by this moral worldview. Early 19th-century researchers such as John Snow, who provided evidence regarding the spread of cholera, sought to apply statistics in the service of social good; that is, of addressing the problems that came alongside the industrial revolution, particularly the spread of disease.

The fourth branch of influence can be traced to Charles Darwin's theory of natural selection, which he famously published in *Origin of the Species* in 1859. Despite Darwin's huge influence in biostatistics, it is important to note that biological applications of statistics had occurred previously for centuries, first as the Catholic Church documented births and deaths in its *registry of souls* and then in the early 19th century, particularly as British bureaucrats sought to determine the *quantum of sickness* through cataloging human existence in *life tables*.

Nevertheless, the late 19th and early 20th centuries underwent a series of social movements focused on manipulating population genetics for the betterment of humanity. This fourth branch can be traced to the ideas of Francis Galton, the developer of one of the first surveys and the creator of *regression* models, who also coined the term *eugenics*, combining the Greek roots for *good* and *genes* or *born*, based on the ideology that improving the genetic composition of society requires both the application of rational thinking and progressive ideas toward understanding and manipulating human population genetics.

In summary, it was a combination of events in the 19th century, predominantly resulting from the *avalanche of numbers* created in the early 19th century, combined with new social problems in during the industrialization of Europe, that led to changes in the way great minds thought about causality and about the role of science. These seeds gestated into what can be called the *golden era* of biostatistics, as it rose to great prominence in human society.

Foundation: 1901–1946

The foundation phase of biostatistics is marked by one of the most remarkable achievements in Western civilization: the adoption across a broad range of disciplines of Darwin's gradualist theory of natural selection. The early 20th-century contributions of members of the *Biometric school* helped resolve a decades-long debate regarding the process of evolution (i.e., whether it is gradual or characterized by a series of *jumps*). The debate was reignited in 1900 with the rediscovery of Gregor Mendel's pea plant experiments. Using quantitative evidence, scientists demonstrated that fundamental questions touching on religion and philosophy, in addition to biology, could be answered scientifically. By the 1940s, scientists reached consensus that Darwin's theory of evolution as a slow, gradual process was empirically valid. This *modern evolutionary synthesis*, as it has been famously called, underscored to both scientists and to the lay public the great possibilities with biostatistics.

The foundation of biostatistics cannot be appreciated without recognition of the discipline's trailblazers in the early 20th century. Members of the Biometric school such as Karl Pearson, R. A. Fisher, and W. F. Weldon achieved fame for viewing evolutionary debate as largely a *statistical problem*, as Weldon put it, and by developing foundations of modern statistical analyses. For example, Fisher developed the *analysis of variance*, the method of *maximum likelihood*, and Pearson developed the eponymous *R correlation coefficient* widely used today. Together, they developed standards for assessing statistical significance and the concept of a *Type I Error* (a false detection of a statistically significant result). The growing methodological development of statistics

as applied to biological questions was marked by the establishment by Fisher of the journal *Biometrika* in 1901.

In addition to the remarkable achievements of the founders of biostatistics, it is important to underscore that they also viewed their developments within the political project of progressivism and of eugenics. Accordingly, Fisher participated in several eugenics societies and published his research in the journal *Eugenics Review*. Largely influenced by Darwin's theory of evolution, eugenicists sought to improve population health; not surprisingly, these ideas transmigrated to Europe, where they were carried into action by the Nazis. After World War II, eugenics as a social movement largely collapsed. Nevertheless, the rise of biostatistics, particularly in American society, maintains at least an implicit Progressivist approach toward societal health, leaving behind eugenics ideology.

Despite the early association of biostatistics with political controversy, its utility as a practice far outweighed any controversy; it is important to note that statistical methods were also developed during World War II, even if these did not have biological applications. Later biostatisticians would build upon the theories developed by Fisher and Pearson, focusing on broader progressive ideals addressing public health more generally.

Institutionalization: 1947 to the Present

Biostatistics did not become its own professionalized field until the mid-20th century. For example, biostatistics first became an independent discipline recognized by the National Institutes of Health in 1946 alongside the establishment of the Center for Disease Control. Professional societies for biostatistics also lagged behind those of statistics. For example, the American Statistical Association was established in 1839, and in England the Royal Statistical Society was established in 1834. The International Biometric Society, however, was not established until 1947, and the Biometrics Section of the American Statistical Association was not adopted until 1950. Other leading journals such as *Statistics in Medicine* were not established until much later, in 1982.

One way of understanding the development and professionalization of biostatistics in the post-war United States is to understand biostatistics

within the *medicalization of society* in the United States more broadly. The term *medicalization*, typically used by sociologists at least since the 1970s, refers to the increasing use of medical language and diagnoses to understand social phenomena. As the 20th century progressed, biostatisticians increasingly influenced the practice of care, as medical professionals used statistics to make important decisions regarding social policy and diagnoses.

A number of advancements occurred within the postwar period, building primarily from the early work of founders. First, physicians and statisticians worked in unison to define and to understand population prevalence of health risks with the development of large-scale surveys. New methods such as survival models were developed to assess mortality and other health risks. Second, with the aid of institutions such as the National Institutes of Health, biostatisticians began developing formal clinical trials to test the efficacy and safety of drugs and other health interventions, largely influenced by R. A. Fisher's theories regarding randomization, but also influenced by the notable statistician and epidemiologist Sir Austin Bradford Hill, who established a set of criteria (the so-called Bradford Hill criteria) for assessing causality in randomized-controlled trials.

In more recent decades, as sociologist Peter Conrad points out, *medicalization* has become increasingly privatized, with biotechnological revolutions, the rise of large pharmaceutical industries, and the rise of managed care. In contemporary Europe and the United States, biostatisticians are increasingly working in commercial fields rather than in academic domains, as private entities rely on biostatistical knowledge to define and simultaneously treat medical problems.

As higher education expanded in the United States and beyond, biostatistics departments began to develop in the postwar era, producing specific credential programs. Nevertheless, the disciplinary boundaries of biostatistics within the academy are still porous at best: the field is often subsumed under statistics departments or amalgamated within schools of medicine, public health, or epidemiology. Nevertheless, increasing numbers of advanced degrees in biostatistics are offered in major universities around the world, almost all of which are postgraduate degrees.

Conclusion: Are We All Biostatisticians Now?

This entry has distinguished biostatistics as a practice and as a profession. As a practice, biostatistics involves the marriage of quantitative methods to biological subject matter. As a profession, biostatistics has grown enormously over the 20th century and thereafter, and, like any profession, has its own institutions, educational credentials, professional groups and societies, and specialty journals. There are two new dynamics that affect the future of biostatistics. In what follows, we will revisit the practice/profession distinction and assess the implications of the big data revolution, so-called, on the behavioral, ecological, educational, and social dimensions of biostatistics.

The *big data revolution* is a term coined in recent years to refer to four aspects of quantitative data analysis. First, owing to increases in computing power, the digitization of records, and the incorporation of technology into everyday social life, there is greater availability of data than has ever been produced before in human history. Second, statistical software is increasingly accessible, with the availability of free, user-generated programs such as R or Python, the increased efficiency of computing power to produce faster results, and the greater availability of more user-friendly interfaces. Third, the big data revolution has led to the availability of online learning sources that have made esoteric statistical knowledge more accessible both within and outside the academy. Fourth, there is a recent cultural revolution: motivation to learn statistics has exploded, as statistical applications have found their way into every aspect of social and professional life. In short, technological advancements have made data more available, analyses easier to implement, skills more accessible, and motivation more widespread.

These implications are both positive and negative for biostatistics. From a practice perspective, biostatistics may far surpass the *golden age* of the 20th century. Soon, nearly everyone may become a biostatistician. From a professional orientation, however, biostatistics may be difficult to maintain as an exclusive profession as more nonprofessionals make substantial contributions to the field. Another concern iterated in biostatistical journals

is ensuring *good* statistical standards in light of greater numbers of participants.

Notwithstanding, if professionalism is the primary bulwark against expanding the contributions of biostatistics, one should be reminded of the pioneers *avant la lettre* who made major contributions even as amateurs. Take for example, John Graunt, often cited in biostatistics as a founding figure, who used public records in an attempt to quell the spread of bubonic plague in 17th-century London. Along the way, he was the first to calculate the population of the city and has been called the *father of demography*. Graunt was, however, a haberdasher by trade. How many future John Graunts are there today, working to contribute to our collective knowledge of the world around us?

Nathan E. Fosse

See also Bioinformatics; Biology, Evolutionary; Health Care Science; Statistics

Further Readings

- Allen, G. E. (2001). Eugenics as an international movement. In N. Smelser & P. Baltes (Eds.), *International encyclopedia of the social & behavioral sciences* (Vol. 16, pp. 4882–4889). Elsevier.
- Borges, J. L. (1975). *A universal history of infamy*. Penguin.
- Bourdieu, P., Wacquant, L., & Farage, S. (1994). Rethinking the state: Genesis and structure of the bureaucratic field. *Sociological Theory*, 12(1), 1–18.
- Daniel, W. W. (2005). *Biostatistics: A foundation for analysis in the health sciences*. Wiley.
- Darwin, C. (1859). *On the origin of species by means of natural selection, or the preservation of favoured races in the struggle for life* (1st ed.). W. Clowes and Sons.
- Dunglison, R. (1865). *Medical lexicon. A new dictionary of medical science, containing a concise account of the various subjects and terms, and formulae for preparations etc* (3rd ed.). Blanchard and Lea.
- Huxley, J. (1943). *Evolution: The modern synthesis*. Harper & Brothers.
- Lew, R. (2005). Experimental design. In P. Armitage & T. Colton (Eds.), *Encyclopedia of biostatistics* (3rd ed., pp. 1835–1844). Wiley.
- Morgan, P. P. (1986). What are my chances of understanding biostatistics? *CMAJ: Canadian Medical Association Journal*, 134(10), 1105–1106.



CATEGORY THEORY

Categories were first introduced by Samuel Eilenberg and Saunders Mac Lane in a paper entitled “General Theory of Natural Equivalences,” published in 1945. A term dating from that time, *general abstract nonsense*, refers to its high level of abstraction compared to classical branches of mathematics. Category theory looks at mathematics on a large scale, examining objects and the relations between them in the abstract, with the aim of stripping away inessential details to get to the essence of things.

This entry provides a brief overview of the history of category theory, and some illustrations of the utility that mathematicians have found in it. It is interesting as an example of how a transformation in how one thinks about the nature of theories at a very high level of abstraction in some discipline can have important implications for how theories are structured and used in that discipline.

Just as objects tend to blur together when we screw up our eyes, category theory allows us to see the underlying similarity between objects previously considered quite different. It provides a way to explore the relationship between different kinds of mathematical objects, for example, products and functions, and abstract from unnecessary details to obtain general definitions and results. It focuses on how things behave, rather on what their internal details are. Category theory is a powerful tool for both philosophers and logicians.

Category theory works from the outside to infer general laws of algebra. An object in category theory is a complete algebra, like a *black box* where we cannot see directly inside and can only know its external properties. We cannot make arbitrary choices about what may be inside an object, we can only describe it as an *isomorphism*. For example, in set theory we think of sets as defined by their members, and we consider functions mapping one set to another in terms of those objects. Category theory, by contrast, treats the mappings between categorical objects directly and does not consider the nature of the internal elements.

However, we can still learn a lot about a category in this way, by investigating how high-level rules, like the axiom of association, are implemented and how identity (e) functions in the category. By doing this in a very general way, we can apply what we learn about one category to other categories.

Much of modern mathematics, from algebraic geometry and homological algebra to logic and topology cannot be imagined without the organizing principles of category theory. It allows us to perform proofs in a purely equational or diagrammatic style of reasoning. It is useful in computing because algebraic structure naturally arises in many different areas of computing, including formal semantics, language design, programming logic, type systems, and programming calculi. Many divisions of the sciences (physics in particular) are discovering categorical ways of thinking to be useful.

History of Category Theory

Category was a term used in Aristotelian mathematics, much later redefined for the purpose of category theory. Categories, functors, and natural transformations were terms first introduced in 1945 when Eilenberg and Mac Lane published a paper titled “The General Theory of Natural Equivalences.”

In 1956, Henri Cartan and Eilenberg published “Homological Algebra,” which enabled mathematicians to learn about algebraic topology and homological algebra in the new categorical language and to understand the new types of diagrams.

In 1957, the German-born French mathematician Alexander Grothendieck published a landmark paper “*Sur quelques points d’algebre homologique*” (“On some points of homologic algebra”) where categories were used intrinsically to define and construct more general theories including algebraic geometry. By using the axiomatic method and language of categories, Grothendieck formulated an abstract type of categories.

In 1958, Daniel M. Kan illustrated that adjoint functors can capture basic concepts in other areas including homotopy theory. Later, category theorists Peter Freyd and Bill Lawvere came to understand that the existence of adjoints to given functors allows there to be definitions of abstract categories.

By the early 1970s, the concept of adjoint functors was a central one to category theory. Category theory became an autonomous area of research, and pure category theory was developed. The theory grew quickly both as a discipline and in terms of its applications to algebraic topology, homological algebra, and algebraic geometry. Lawvere proposed the *theory of categories* for set theory, and thus the entire realm of mathematics, and used categories for the study of logical aspects of mathematics. He axiomatized the category of sets, gave a categorical description of theories independent of syntactical choices, and argued that functors should play a foundational role in category theory.

Present-day category theory is used in theoretical computer science, mathematical physics, and has contributed to the development of new logical systems.

Categories, Objects, and Morphisms

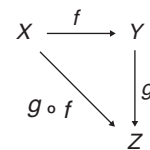
In everyday speech, we think of a category as a kind of thing. A category consists of a collection of things all of which are related in some way. In mathematics, a category can also be looked at as a collection of *things*, and a type of relationship between pairs of such things.

To count a duo of things as a category we need to adhere to two rules: That every *thing* be related to itself by simply being itself, and if one thing is related to another and the second is related to a third, then the first is related to the third. In a category, *things* are called *objects* or *sets*, and the relationships or typed total functions are called *morphisms*.

Morphisms are the mappings that preserve the structures between objects. By examining these morphisms and the way they compose or their paths commute, we can understand more about the object’s structure.

Relations between morphisms are often depicted in commutative diagrams with *points* or corners, representing objects and *arrows* representing morphisms. By studying the arrows, we can look at properties of spaces that do not involve drawing complicated pictures in higher dimensions, since the space is represented by a single point and arrows are one-dimensional.

For example,



The existence of an identity arrow for each object and being able to compose the arrows associatively are basic properties of a category. A simple example is the category of sets, where the objects are sets and the arrows are functions.

However, the objects of a category do not have to be sets and the arrows do not need to be functions; any method of formulating a mathematical concept to meet the basic conditions of the object’s and arrows’ behavior is a valid category, so that category theory will apply to it.

Categories are considered identical if they have the same collection of objects and arrows and use the same associative method to align pairs of arrows.

The class $\mathbf{Grp}$ of groups consists of all objects having a *group structure*. We can prove theorems about groups by making logical deductions from sets of axioms. For example, we can immediately prove from the set of axioms that the group's identity element is unique.

When discussing groups, these morphisms are termed *homomorphisms*. A group homomorphism between two groups *maintains the group structure* precisely—it is a method of taking one group to another, so that information about the structure of the first group gets carried to the second group. The study of group homomorphisms then provides a tool for studying general properties of groups and consequences of the group axioms.

Well-known categories are denoted by a shortened capitalized word or abbreviations in bold or italics: examples are **Sets**, the category of sets and set functions; **Rings**, the category of rings and ring homomorphisms; and **Tops**, the category of topological spaces and continuous maps. All the former categories use the identity map as an identity arrow and composition as the associative process on the arrows.

Functors

A category is a type of mathematical structure, so we can look for processes that preserve this structure. This process is known as a *functor*. Functors are represented by arrows between categories. They can define categorical diagrams and sequences and associate to every object of a category an object of another category and to every morphism in the first category a morphism in the second.

Objects are categories, and morphisms between categories are functors. By studying categories and functors, we are not simply studying a class of mathematical structures and the morphisms between them, we are studying the relationship *between various classes of mathematical structures*. Difficult topological questions can be translated into algebraic questions which make solving them easier. Basic constructions, for example, the fundamental group or fundamental groupoid of a topological space, can be expressed as fundamental functors to the category of groupoids in this way, and the concept is pervasive in algebra and its applications.

Natural Transformations

Saunders Mac Lane, one of the founders of category theory, stated that his objective was not to study functors but to study natural transformations. Without a study of functors, the study of categories would not be complete. But the study of functors could not be complete without the study of natural transformations. This needs to be put in the context of the axiomatic theory of *homology*. Homology in algebraic topology and abstract algebra is a way to associate a sequence of abelian groups within a given mathematical object, for example, a topological space or group.

Some diagrammatic and some sequential constructions can be *naturally related*—which can at first glance seem vague. This leads us to the clarifying concept of natural transformation, which is a way to map one functor to another. This has led to a context in which many important mathematical constructions can be studied. *Naturality* is a principle in physics that is more significant than it first appears. An arrow or morphism between two functors is called a *natural transformation* when certain conditions such as naturality or commutativity apply to it.

Functors and transformations are the key concepts in category theory. A natural transformation can transform one functor into another while maintaining its internal structure that is the composition of morphisms in the category. A natural transformation is essentially a morphism of functors, an intuition that can be used to define functor categories. Natural transformations are fundamental to category theory and are evident in most of its applications. Natural transformations are categorical in nature, so particular maps between functors are consistent throughout a category.

A particular map between individual objects, rather than entire categories, is termed a *natural isomorphism*. This means that it is defined on the whole category and defines a natural transformation of functors. There is also such a thing as an *unnatural isomorphism*, which is created if a category-wide map cannot be extended to a natural transformation.

Equivalence

Equivalence is a relationship between two objects in a similar way to equality and isomorphism but

is weaker and more general. Isomorphism is weaker than equality and equivalence is weaker still. If we take two mathematical objects, A and B, to be compared and we construct two morphisms between them, C and D, then E and F must be selected so that by applying them (composing them) we get back to the isomorphic from whence we started.

Category theory is a broad topic and has the potential to be applied across many branches of mathematics. From a philosophical point of view, category theory provides a challenge to philosophers to tackle the general, ontological, and epistemological status of categories and categorical methods.

Thora L. Fitzpatrick

See also Axiomatic Theory; Mathematics, 19th Century; Mathematics, 20th Century; Realism in Mathematics; Set Theory

Further Readings

- Asperti, A. (1991). *Categories, types and structure: Introduction to category theory for the working computer scientist*. MIT Press.
- Awodey, S. (2010). *Category theory* (2nd ed.). Oxford University Press.
- Herrlich, H., & Strecker, G. E. (1979). *Category theory: An introduction*. Heldermann Verlag.
- Lawvere, F. W., & Schanuel, S. H. (1997). *Conceptual mathematics: A first introduction to categories*. Cambridge University Press.
- Mac Lane, S. (1998). *Categories for the working mathematician*. Springer.
- Simmons, H. (2011). *An introduction to category theory*. Cambridge University Press.

CELL THEORY

The cell is the basic unit of organization or structure of living matter. All animals and plants are composed of cells and cell products; all tissues and organs are composed of cells. Among the various landmark biological theories formulated in the 19th century, cell theory was to become a great unifying concept and one of the key principles of biology. Indeed, in explaining the

organization of living systems, cell theory provides the foundation for understanding the growth and reproduction of organisms. No matter how big or small, or how simple or complex they might be. This entry takes a close look at the components and characteristics of cells, plant, and animal, as defined by cell theory and its close relative, biogenesis theory.

Origins and Basic Principles

It was the invention and refinement of the microscope by Dutch lens makers Hans and Zacharias Janssen in the 17th century and subsequent observations of tiny specimens by the Englishman Robert Hooke and the Dutch self-taught scientist Anton van Leeuwenhoek—who was the first to observe bacteria and other single-celled organisms—that eventually gave rise to cell theory and its cousin, biogenesis theory. Cell theory, first proposed in the late 1830s by two German scientists, Matthias Jakob Schleiden (a botanist) and Theodor Schwann (a zoologist), posits that—as affirmed by more than a century of published reports by trained observers—all living things are composed of cells. Biogenesis theory, proposed in 1858 by the German physician, historian, and pathologist Rudolf Virchow, states that all living cells arise from existing (living) cells, and that no cells are created spontaneously from nonliving matter. Classical cell theory, as developed through the observations of Hooke, Leeuwenhoek, Schleiden, Schwann, Virchow, and others as a fundamental theory in biology, includes the following three principles:

1. All organisms (animals, plants, and microbes) are composed of one or more cells.
2. The cell is the smallest unit having the properties of life, the basic unit of organization of living organisms (i.e., the cell is the basic unit of life).
3. All cells come from preexisting cells; that is, no cell can originate spontaneously but comes into being only by division and duplication of already existing cells.

These basic principles or ideas still hold true and they provide the foundation for everything that modern researchers have learned about cells.

Cell theory has had wide biological applications. With the progress of biochemistry, it was shown that there were fundamental similarities in the chemical composition and metabolic activities of all cells.

Discoveries That Led to Cell Theory

As noted earlier in this section, cell theory was developed from three German scientists' discoveries. They are Matthias Schleiden, Theodor Schwann, and Rudolph Virchow. In 1838, Schleiden discovered that all plant tissues he observed under the microscope were composed of cells. Then only a year later (1839), Schwann discovered that all animal tissues he observed in similar fashion were composed of cells. Later, in 1855, a Virchow was doing experiments with diseases when he found that all cells come from other existing cells. Cells of course were discovered much earlier. The first person to see a cell was the English scientist Robert Hooke (1635–1703). He used a very primitive microscope, but when he was using it to examine a thin slice of cork, he saw cells for the first time. Their shape reminded him of monks' cells in monasteries and so he dubbed them cells. As mentioned above, the first person to see living cells was Anton Van Leeuwenhoek, a Dutch microscope builder.

Modern Cell Theory

The modern tenets of cell theory include the following:

- All known living things are made up of cells.
- The cell is the fundamental unit of structure and function in living things.
- All cells come from preexisting cells by division.
- Cells contain hereditary information (DNA) which is passed from cell to cell during cell division.
- All cells are basically the same in chemical composition.
- All energy flow (metabolism and biochemistry) of life occurs within cells.
- Some organisms are unicellular, made up of only one cell.
- Other organisms are multicellular, composed of countless number of cells.

- The activity of an organism depends on the total activity of independent cells.
- As with any theory, its tenets are based upon previous observations and facts, which are synthesized into a coherent whole via the scientific method.

Cell Organization and Size

All living cells have three things in common. They have an outer plasma membrane, deoxyribonucleic acid (DNA), and cytoplasm. The plasma membrane is a structurally distinctive surface membrane that surrounds all cells. Every cell is enveloped by a protective plasma membrane, which is the outer covering that encloses the cell's internal parts and separates its contents from the surrounding external environment, so that cell activities can go on apart from events that may be taking place outside the cell. Substances can still move across the membrane, which does not completely isolate the cell's interior but serves to regulate passage of materials between cell and environment. Cells exchange materials with the environment and can accumulate needed substances and energy stores. Cells have specialized molecules that contain genetic instructions. In most cells, the genetic instructions are encoded in deoxyribonucleic acid, more known as DNA. The DNA is to be found inside the cell, along with molecules that can copy or read the inherited genetic instructions DNA carries. A semifluid matrix called *cytoplasm* fills the interior of the cell. The cytoplasm contains all of the sugars, amino acids, and proteins the cell uses to carry out its everyday activities. Although it is aqueous medium, cytoplasm is more viscous than water, owing to the high concentration of proteins and other macromolecules. It consists of a thick, jelly-like fluid, the cytosol, and various other components. Typically, cells have internal structures, called *organelles*, that are specialized to carry out metabolic activities such as converting energy to usable forms, synthesizing needed compounds, and manufacturing structures necessary for functioning and reproduction.

Although most cells are small, some quite large cells do exist. Their sizes vary over a wide range. A few cells, including the yolk of chicken eggs, can be seen with the unaided eye but most cells are so

small that they can only be seen with a microscope and must be measured in very small units. The surface-to-volume ratio is responsible for the small size of cells. As a cell's size increases, its volume increases much more rapidly than its surface area. The ratio is a physical relationship. It dictates that as the linear dimensions of a three-dimensional object increase, the volume of the object increases faster than its surface area does. For a spherical cell, surface area is proportional to the square of the radius, whereas the volume is proportional to the cube of the radius. Thus, if the radii of two cells differ by a factor of 10, the larger cell will have 10^2 , or 100 times, the surface area, but 10^3 , or 1,000 times, the volume of the smaller cell. In another example, if a round cell grew like an inflating balloon so that its diameter increased to four times the starting girth, the volume inside the cell would be 64 times more than before, but the cell's surface would be just 16 times larger. The cell would not have enough surface area to allow nutrients to flow inward rapidly, or for wastes or cell products to move rapidly outward.

The sizes and shapes of cells are related to the functions they perform. Some cells—such as the protozoan called the *amoeba*, and white blood cells—change their shape as they move about. Sperm cells, male reproductive cells, have long, whiplike tails, called *flagella*, for locomotion. Nerve cells have long, thin extensions that enable them to transmit messages over great distances. The extensions of some nerve cells in the human body may be as long as one meter. Other cells, such as certain epithelial cells, such as those that form the skin and line the intestines, are almost rectangular and are stacked much like building blocks to form sheetlike structures.

Cell Function

Cells organize living things. They keep chemical processes tidy and compartmentalized so individual cell processes do not interfere with others and the cell can go about its business of metabolizing, reproducing, and so on. As noted earlier, cell components are enclosed in a membrane which serves as a barrier between the outside world and the cell's internal chemistry. The cell membrane is a selective barrier, meaning that it lets some

chemicals in and others out, and in doing so maintains the balance necessary for the cell to live. The cell surface provides the only opportunity for interaction with the environment, because all substances enter and exit a cell via this surface. The membrane surrounding the cell plays a key role in controlling cell function.

Cell Types

All living organisms can be sorted into one of two groups depending on the fundamental structure of their cells. These two groups are the prokaryotes (prokaryotic cells) and the eukaryotes (eukaryotic cells). Prokaryotic cells are typically smaller than eukaryotic cells. In fact, the average prokaryotic cell is only about one-tenth the diameter of the average eukaryotic cell.

Prokaryotic Cells

Prokaryotes are organisms made up of cells that lack a cell nucleus or any membrane-enclosed organelles. Prokaryotic cells (prokaryotic means *before the nucleus*) are small cells that lack complex interior organization. They consist of cytoplasm surrounded by a plasma membrane that confines the contents of the cell to an internal compartment, with a characteristic cell wall that may differ in composition depending on the particular organism. In some prokaryotic cells, the plasma membrane may be folded inward to form a complex of membranes along which many of the cell's metabolic reactions take place. The plasma membrane of a prokaryotic cell carries out some of the functions organelles perform in eukaryotic cells. Nothing separates the cell's DNA from other internal cell parts. Prokaryotes have no distinct interior compartments; they lack a nucleus (although they do have circular or linear DNA) and other membrane-bound organelles (though they do contain ribosomes). Many prokaryotes have flagella, long fibers that project from the surface of the cell. Prokaryotic flagella, which operate like propellers, are important in locomotion. The dense internal material of the bacteria cell contains ribosomes, small complexes of ribonucleic acid (RNA), and protein that synthesize polypeptides, most lack the membrane-bounded

organelles characteristic of eukaryotic cells. The ribosomes of prokaryotic cells are smaller than those found in eukaryotic cells. Prokaryotic cells also contain storage granules that hold glycogen, lipid, or phosphate compounds. Bacteria and a microscopic organism called *archaea* are the only prokaryotic cells. A typical anatomy of prokaryotic cell might contain the following parts: cell wall, plasma membrane, cytoplasm, flagella and pili, nucleoid, and plasmid.

Eukaryotic Cells

Eukaryotes are organisms made up of cells that possess a membrane-bound nucleus (that holds genetic material) as well as membrane-bound organelles. The eukaryotic (*true nucleus*)

organisms have cells with membrane-bound nuclei containing chromosomes composed of chromatin. They occur in plants (from algae to angiosperms) and animals (from protozoa to mammals). The hallmark of the eukaryotic cells is compartmentalization. This is achieved through a combination of an extensive endomembrane system that weaves through the cell interior and by numerous organelles. In their cytoplasm are tiny compartments and sacs called *organelles* (*little organs*). One organelle, the nucleus, contains the DNA of a eukaryotic cell that carries out specialized functions. The nucleus, composed of a double membrane connected to the endomembrane system, contains the cell's genetic information. The eukaryotic cell has its own control center, internal

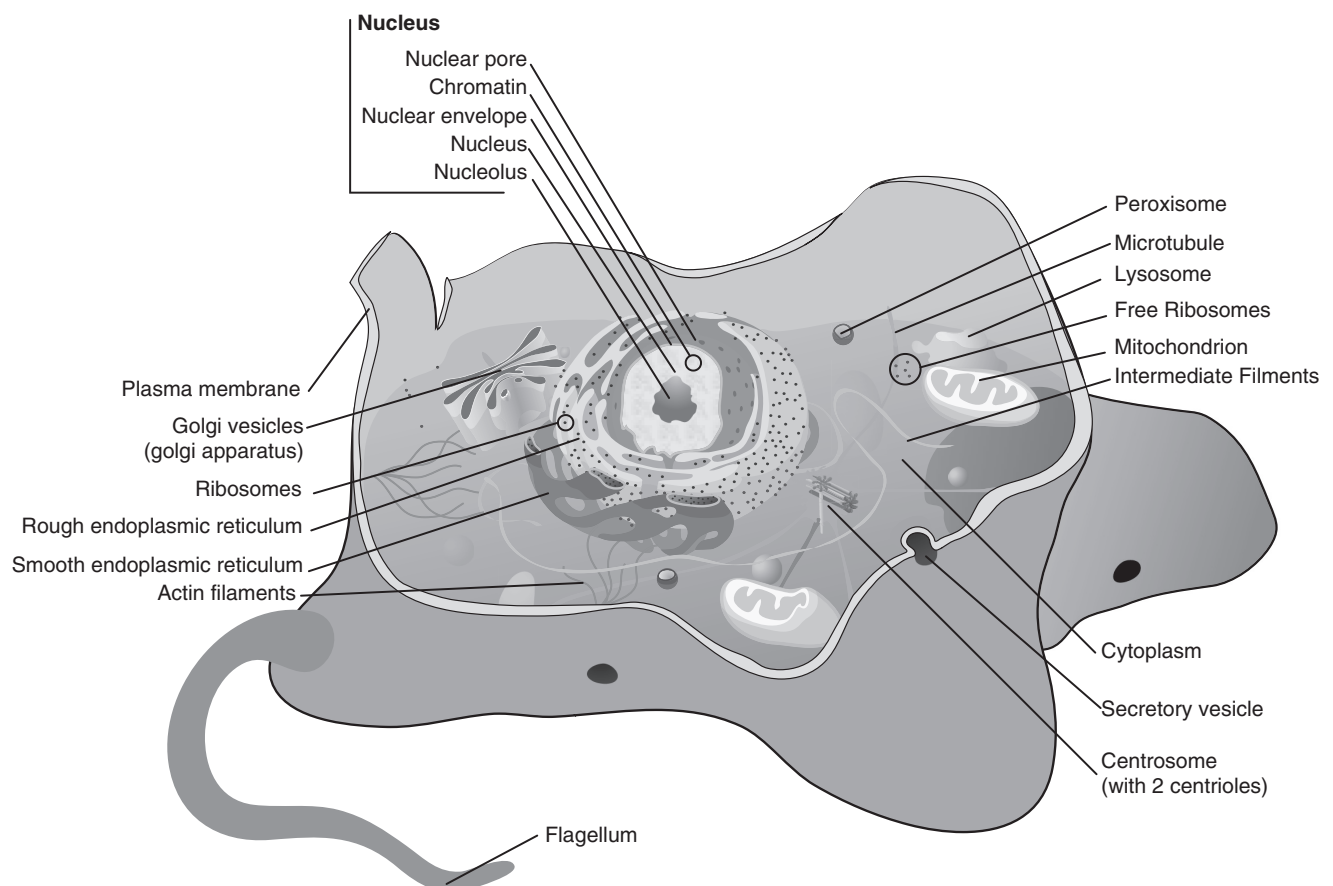


Figure 1 Structure of an Animal Cell

transportation system, power plants, factories for making needed materials, packaging plants, and even a *self-destruct* system. Biologists refer to the part of the cell outside the nucleus as *cytoplasm* and the part of the cell within the nucleus as *nucleoplasm*. Various organelles are suspended within the fluid component of the cytoplasm, which is called the *cytosol*. Therefore, the term *cytoplasm* includes both the cytosol and all the organelles other than the nucleus. The many specialized organelles of eukaryotic cells solve some of the problems associated with large size, so eukaryotic cells can be much larger than prokaryotic cells. Eukaryotic cells also differ from prokaryotic cells in having a supporting framework, or cytoskeleton, important in maintaining shape and transporting materials within the cell. Though the eukaryotic cells have different shape, size, and physiology; all the cells are typically composed of the following parts: cell wall, plasma membrane, cytoplasm, nucleus, nucleolus, chromosomes, ribosome, vesicle, endoplasmic reticulum (ER), Golgi apparatus, cytoskeleton, lysosome, centrioles, and mitochondria.

The Structure of Cells

Cell Wall

The outermost structure of most plant cells is a rigid layer called the *cell wall*. It is composed mainly of carbohydrates such as cellulose, pectin, hemicelluloses and lignin, and fatty substances such as waxes. Ultrastructurally, the cell wall is found to consist of a microfibrillary network within in a gel-like matrix. The microfibrils are mostly made up of cellulose. There is a pectin-rich cementing substance between the walls of adjacent cells, which is called the *middle lamella*. The primary cell wall is formed immediately after the division of a single cell into two cells. Many kinds of plant cells have only primary cell wall around them. In certain types of plant cells, such as those that constitute the tissues phloem and xylem, an additional layer is added to the inner surface of the primary cell wall at a later stage. Called the *secondary cell wall*, it consists mainly of cellulose, hemicellulose, and lignin. In many plant cells, there are tunnels running

through the cell wall called *plasmodesmata*, which allow communication with the other cells in a tissue. The cell wall constitutes a kind of exoskeleton that provides protection and mechanical support to the plant cell. It determines the shape of plant cell and prevents it from desiccation.

Plasma Membrane

Every kind of animal cell is bounded by a living, extremely thin and delicate membrane called the *cell membrane*, *plasmalemma*, or *plasma membrane*. Membranes divide the eukaryotic cell into compartments, and their unique properties enable membranous organelles to carry out a wide variety of functions. For example, cell membranes never have free ends; therefore, a membranous organelle always contains at least one enclosed internal space or compartment. These membrane-enclosed compartments allow certain cell activities to be localized within specific regions of the cell. Compartmentalizing also allows many different activities to go on simultaneously. Membrane-enclosed compartments keep certain reactive compounds away from other parts of the cell that they might adversely affect. Membranes allow cells to store energy. A difference in the concentration of some substance on the two sides of a membrane is a form of stored energy or potential energy. The plasma membrane is a selectively permeable membrane; its main function is to control selectively the entrance and exit of materials. This allows the cell to maintain a constant internal environment or *homeostasis*. The cell membrane regulates the crossing of chemicals in and out of the cell in several ways: by *diffusion* (the tendency of solute molecules to minimize concentration and thus move from an area of higher concentration toward an area of lower concentration until concentrations equalize), *osmosis* (the movement of solvent across a selective boundary in order to equalize the concentration of a solute that is unable to move across the boundary), and *selective transport* (via membrane channels and membrane pumps). As particles of substance move across the membrane from the side of higher concentration to the side of lower concentration, the cell converts some of this potential energy to the chemical energy of *adenosine triphosphate* (ATP)

molecules. This process of energy conversion is a basic mechanism that cells use to capture and convert the energy necessary to sustain life.

Cytoplasm

Within the plasma membrane is the cytoplasm, occupying the space between the plasma membrane and the region of DNA. It consists of a thick gel that surrounds the organelles, the cytosol, and various other components. Cytosol, the matrix of the cytoplasm, constitutes the cell's true internal milieu. The cytosol serves to dissolve or suspend the great variety of small molecules concerned with cellular metabolism such as glucose, amino acids, nucleotides, vitamins, minerals, oxygen, and ions. Cell biologists have identified many types of organelles in the cytoplasm of eukaryotic cells. Among them are the ER, ribosomes, Golgi complex, lysosomes, peroxisomes, vacuoles, mitochondria, and chloroplasts.

ER. The ER is a transport network for molecules targeted for certain modifications and specific destinations, as compared to molecules that float freely in the cytoplasm. The ER has two forms: the rough ER, which has ribosomes on its surface that secrete proteins into the ER, and the smooth ER, which lacks ribosomes. The smooth ER plays a role in calcium sequestration and release.

Ribosome. The ribosome is a large complex of RNA and protein molecules. Each consists of two subunits and acts as an assembly line where RNA from the nucleus is used to synthesize proteins from amino acids. Ribosomes can be found either floating freely or bound to a membrane (the rough ER in eukaryotes or the cell membrane in prokaryotes).

Golgi apparatus. The primary function of the Golgi apparatus is to process and package the macromolecules such as proteins and lipids that are synthesized by the cell.

Lysosomes and peroxisomes. Lysosomes contain digestive enzymes (acid hydrolases). They digest excess or worn-out organelles, food particles, and engulfed viruses or bacteria. Peroxisomes have enzymes that rid the cell of toxic peroxides. They are involved in lipid metabolism and detoxify harmful compounds.

Vacuoles. Vacuoles store food and waste. Some vacuoles store extra water. They are often described as liquid-filled space and are surrounded by a membrane.

Some cells, most notably the amoeba, a protozoan, have contractile vacuoles, which can pump water out of the cell if there is too much water. The vacuoles of eukaryotic cells are usually larger in those of plants than of animals.

Mitochondria and chloroplasts. Mitochondria are self-replicating organelles that occur in various numbers, shapes, and sizes in the cytoplasm of all eukaryotic cells. Mitochondria play a critical role in generating energy in the eukaryotic cell. Respiration occurs in the cell mitochondria, which generate the cell's energy by oxidative phosphorylation, using oxygen to release energy stored in cellular nutrients to generate ATP. Chloroplasts can only be found in plants and algae, and they capture the sun's energy to make ATP through photosynthesis.

Nucleus

The nucleus is the most prominent organelle in the cell. The nucleus is often described as a cell's master control center. It is also a protective *isolation chamber* for DNA, the genetic material. Nucleus is usually spherical or oval in shape and centrally located component which controls all the vital activities of the cytoplasm and carries the hereditary material the DNA in it. Because of its size and the fact that it often occupies a relatively fixed position near the center of the cell, some early investigators guessed long before experimental evidence was available that the nucleus served as the control center of the cell. The nucleus encloses the DNA of a eukaryotic cell. DNA contains instructions for building a cell's proteins. Those proteins in turn determine a cell's structure and function. In a human cell, there are 46 DNA molecules that together would be more than six feet long if they were stretched out end to end. Most cells have one nucleus, although there are exceptions.

The basic structure of the nucleus is composed of the

Nuclear envelope, which consists of two concentric nuclear membranes (an inner nuclear membrane which is lined by nuclear lamina and an outer nuclear membrane which is continuous with ER) that separate the nuclear contents from the surrounding cytoplasm. The envelope surrounds the fluid part of the nucleus (the nucleoplasm), and many proteins are embedded in its layers.

Nuclear pores, which consist of protein complexes, regulate the passage of materials between nucleoplasm and cytoplasm.

Nucleoplasm, the fluid interior portion of the nucleus.

Chromatin, which includes all the DNA molecules and their attached proteins; it appears as a network of granules and strands in cells that are not dividing. The chromatin has two forms: euchromatin and heterochromatin. Euchromatin is the well-dispersed form of chromatin which takes lighter DNA stain and is genetically active, that is, it is involved in gene duplication, gene transcription, and phenogenesis or phenotypic expression of a gene through some type of protein synthesis. Heterochromatin is the highly condensed form of chromatin which takes dark DNA stain and is genetically inert.

Chromosomes, composed of the individual DNA molecules and their attached proteins. *Nucleolus* which is a cluster of the RNA and proteins used to assemble ribosome subunits. It is the site where ribosomes are manufactured.

Cytoskeleton

Scientists watching cells growing in the laboratory see that cells frequently change shape and that many types of cells move about. The *cytoskeleton* is a dynamic internal framework made of microtubules, microfilaments, and intermediate filaments. The cytoskeleton provides structural support and functions in various types of cell movement, including transport of materials in the cell.

The Question of Viruses

Although there is no precise definition of what separates the living from the nonliving, it is widely accepted that because viruses consist not of cells, but only of DNA and RNA surrounded by a protein shell, they are not living organisms but complex biological mechanisms. Viruses lack a plasma membrane and metabolic machinery for energy production and for the synthesis of proteins. In consequence, viruses can only reproduce inside a suitable host cell, which may belong to animals, plants, or bacteria. Thus, a virus may be defined as an infectious, subcellular, and ultramicroscopic particle representing an obligate cellular parasite

and a potential pathogen whose reproduction (replication) in the host cell and transmission by infection cause characteristic reaction in the host cells. Outside the host cells, viruses are just like nonliving inert particles and, like salt or sugar, they can be purified, crystallized, and placed into jars on a shelf for years. They use their own genetic program for reproduction but rely on the raw materials (i.e., amino acids and nucleotides) and biosynthetic machinery of the host cells for their multiplication.

Samson Olowo Kolawole

See also Developmental Systems Theory; Germ Theory

Further Readings

- Cooper, G. M. (2000). *The cell: A molecular approach* (2nd ed.). Sinauer Associates, Inc.
- Gilbert, S. F. (2000). *Developmental biology* (6th ed.). Sinauer Associates, Inc.
- Hickman, C. P., Jr., Roberts, L. S., Larson, A., & P'Anson, H. (2004). *Integrated principles of zoology*. McGraw-Hill.
- Kalthoff, K. (2001). *Analysis of biological development* (2nd ed.). McGraw-Hill.
- Plopper, G., & Ivankovic, D. B. (2021). *Principles of cell biology* (3rd ed.). Jones & Bartlett Learning.
- Raven, P., Johnson, G., Mason, K., Losos, J., & Duncan, T. (2020). *Biology* (12th ed.). McGraw-Hill.
- Rost, T. L., Barbour, M. G., Stocking, C. R., & Murphy, T. M. (2006). *Plant biology* (2nd ed.). Thomson Brooks/Cole.
- Solomon, E. P., Berg, L. R., & Martin, D. W. (2006). *Biology* (7th ed.). Brooks/Cole-Thomson Learning.
- Starr, C., & McMillan, B. (2010). *Human biology* (8th ed.). Brooks/Cole Cengage Learning.
- Verma, P. S., & Agarwal, V. K. (2010). *Cell biology, genetics, molecular biology, evolution and ecology*. S. Chand & Company Ltd

CHEMISTRY

Chemistry is understood as the set of scientific and technological practices based on analysis and synthesis that transform substances. Many of these substances have been and are of enormous

social importance, for example, when they are marketed as new materials or as medicines that improve and prolong human life.

As we know it today, chemistry is the result of the multiple legacies of a thousand years that, embodied in trades, influenced the daily life of all cultures, helping to build everywhere a material culture. It is still surprising that practices as different as that of the blacksmith and metallurgy, the healer and pharmacy, the potter and ceramics, and the baker and biotechnology have come together to end up merging, barely three centuries ago, in a common field: chemistry, by which practice substances are analyzed and transformed, in both small and large quantities, and over short or long spans of time. The following sections of this entry examine the origins and advances of theories and practices in chemistry.

Beginnings

It should be noted at the outset that analysis and synthesis of substances have been foundational for chemistry.

The analysis of substances, long associated with the concept of purity, has been a primary concern, if not an obsession, for chemists. Because substances found in nature cannot be considered pure, the separation of their constituent parts, the isolation of what is wanted, has been a constant in chemical practices, from their beginnings in the ancient tradition of alchemy, which spans four millennia and three continents, up to the present day. Among contemporary scientists, it is clear that there is no such thing as pure substances. Instead, there is a model of pure substance that has been built at the interface of technoscience over the years. What we have direct access to is a *predominant* substance mixed in minor or very minor amounts with other, different substances. Purity depends on our technical ability to identify impurities. Often, the modifications in chemical practices are the result of experimental innovations, associated with the incorporation of new instruments, rather than by the introduction of a new theory.

Through the operation of technical-chemical systems, human beings, as willing agents, have obtained substances that do not occur in nature, such as dynamite, aspirin, nylon, chlorofluoroalkanes, and the millions of artificial substances

that constitute a supnature, and which are philosophically called *artifacts*. There are no new substances—or artifacts—without action and without design. They are not only the result of an intentional human action, they also have a meaning embedded in a specific historical context. Since its millennial origin, through the trades, the main way in which chemists today *know* is by *doing*, and this chemical practice characterized by action has always increased the world's complexity. Chemistry professionals (a few million people around the world) make fundamentally new substances. Through their synthesizing practices, the number of such substances grew from several hundred in 1800 to more than 150 million at the beginning of the 21st century, most of which are on the market. And more than 15,000 new ones are being added every day, that is to say, one every 6 seconds. The synthesis of new substances makes chemistry an extraordinarily productive science. *Chemical Abstracts*, the database that informs us about most of the publications in this discipline, have reported practically the same number of publications as all the other sciences put together. This tremendously successful quality of chemical practices, this proliferation of artifacts, has caused an ethical debate about their impact on the world.

It should be noted that a disciplinary knowledge is always connected to a social context of transmission and to a social group, in this case the chemical community, which reproduces itself. For this reason, what we currently call chemistry, both academic and industrial, can only be understood taking into account its major transformations throughout history.

Chemical Practices: Pluralism, Theories, and Models

Chemistry is, therefore, a relatively young discipline that has integrated a multitude of millenary trades, today transformed into technosciences. Chemistry is about the changes in substances, whether in small or very large amounts. The process through which this happens is called a *chemical reaction*, in which one or more substances, the reactants, change to one or more different substances, the products. The time spans of these

reactions vary: They can occur almost instantaneously or over a prolonged period. Chemistry itself can be recognized as a practice—that is to say, a series of coordinated and shared activities (rules, reasons, techniques, purposes, beliefs) that are disciplined through the change of *correct* norms or procedures, and which have a stable structure that can reproduce itself through various learning processes. Some of the norms and procedures have changed slowly and some others quickly. Chemistry is recognized as a practice because the examination of the reasoning in chemists' ongoing work, particularly in experimental work, rests upon tacit knowledge, such as cannot be expressed in words. As in any industrial business, pragmatism is also required.

The individual ideal agent—the scientist—is replaced by a specific practitioner inside a specific community, for example, a chemical industry within a particular context.

According to some observers, chemistry fits better with models than with theories. Models are built to explain and to predict certain aspects of the world—with a specific goal, related, for example, to a certain chemical practice, as in substitution reactions. Because in chemistry, models are also mediators between the world and ourselves, which means that they function not only as representations but also as means of intervention; for example, Berzelian formulas as paper tools in 18th-century organic chemical practice were originally written H_2O for water or $\text{C}_2\text{H}_6\text{O}$ for ethanol. Different models for the same field of application can coexist and usefully complement each other, for example, in actual acid–base reactions, in contrast to theories that have been characterized as abstract conceptual objects with a logical structure.

In the scientific practice of chemistry, the chemical experiment is carried out in the laboratory. Chemical practices—analysis and synthesis—are different from other scientific practices, particularly those in physics. By participating in a practice, one knows what to do and what to say. Chemical practices are related to the historian of science Thomas Kuhn's *exemplars*, that is to say, the collection of problems, both theoretical and experimental, shared and solved by a specific community at a particular historical moment, which are generally found in professional

publications, academic papers, patents, and especially in a given discipline's textbooks. Apprenticeship in the regimented discipline of the chemical community allows the transmission of purposes and also the possibility of improving the practices themselves. Chemical practices do not try to discover how matter is constituted; what they mainly seek is to build new substances or to improve the social applications of the already known. For chemists, reality is found in the entities that explain chemical practices, such as molecular orbitals, rather than in the underlying physical theories, like quantum mechanics.

In accepting that the history of chemistry can be reconstructed from the practices carried out by different communities, pluralism is required. In other words, the facts emanating from small research laboratories, greatly amplified through industrial processes, related to multiple purposes but limited by the complexity of the chemical systems themselves, demand pluralism.

Pluralism supports that there is more than one way to success. The pluralism of purposes in chemical practices for some practitioners consists in analyzing the composition of a certain plant, whereas for some others, the purpose is to synthesize a new substance similar to the one found in the plant; and for some others to produce it in large quantities or in shorter time; and for still others verifying that it does not pollute the environment; hence, the idea of monism identified with a single scientific method is rejected.

Although Kuhn, in *The Structure of Scientific Revolutions*, treated the chemical revolution (marking the birth of this discipline as an independent science) as a clear example of a paradigm shift, his proposal has since been debated. In his discourse, there is scant discussion on experimental work and the use of instruments, a matter on which philosophers and historians of science have worked intensively in recent years. In a challenge to the Kuhnian picture, this means that during those historical periods that have been characterized as revolutionary, not all previous concepts are abandoned, let alone the experimental procedures, which are transformed from within by shifting the questions and the acceptable answers—after which new subdisciplines have appeared. In other words, following a major historical chemical transformation, different chemical practices are

introduced with new subdisciplines, although some exemplars remain and others change. This continuity and the rupture process are further explained in the next section through a brief history of chemical practices.

Here, it is important to recognize that the theoretical discoveries of chemical communities throughout their history have received different names: theories (e.g., Bronsted–Lowry acid–base theory), equations (e.g., Arrhenius equation), laws (e.g., constant proportions law), principles (e.g., Le Chatelier principle), and models (e.g., Lewis atomic model). Since chemical models allow a more successful interpretation of plural chemical practices, from now on, they will be used in this entry.

A Brief History of Chemical Practices

Chemistry emerged as a science, with robust and widespread laboratory and industrial practices in Europe, in the mid-18th century. Since then, it has undergone four major transformations, each characterized by the emergence of new subdisciplines like organic chemistry, physical chemistry, instrumental chemistry, and organometallic chemistry, among others.

With the waning of alchemy, there was a transitional period known as *protochemistry*. This can be characterized by the construction of at least three models capable of explaining what we know today as chemistry. The mechanistic model of particles developed by Robert Boyle; the compositional model developed by G. E. Stahl, which considered that chemical reactions, particularly those of combustion, were carried out from the presence of a substance called *phlogiston* in everything that contained it; the model of affinities developed by E. F. Geoffroy, which collected the empirical experience by which some substances were more inclined to combine with each other.

Chemistry as we know it today began in 1754 when the highly successful lecturer Joseph Black improved the analytical balance and isolated carbon dioxide from magnesium carbonate, in what can be recognized as the first quantitative chemical reaction. Over a decade later, Henry Cavendish improved the hydropneumatic tank and isolated a new gaseous substance not contained in natural air. Hydrogen was the first gas recognized as such. Antoine. Lavoisier is still considered a

key figure in this period, merging two independent and different chemical traditions, continental analytical chemistry and British pneumatic chemistry. With his definition of an element as the last unit of empirical analysis, he realized his ambition to reform and improve the chemical nomenclature. With the development of his calorimeter, the oxygen model to explain combustion reactions and the determination of the model of conservation of matter (commonly known as *law of conservation of matter*), the way of new chemical practices was cleared.

In 1749, John Roebuck and Samuel Garbett pioneered the industrial production of sulfuric acid by the lead chamber process. Subsequently, sulfuric acid became a very important chemical commodity, and its national production was used as a good indicator of the industrial strength of each country. Nicolas LeBlanc patented the method for the large-scale production of sodium carbonate, obtained from sulfuric acid and sodium chloride. Sodium carbonate was used mainly in the manufacture of soaps, glass, and paper. The production of hundreds of tons of both substances per year meant the beginning of industrial chemistry and its impact on the environment.

Alessandro Volta's invention of the electric battery in 1800 was followed by a multitude of experiments in which some researchers passed electricity through different objects and substances. One of them, Humphrey Davy, stood out for this activity by the isolation of many elements, among them sodium (Na), magnesium (Mg), calcium (Ca), and strontium (Sr). The birth of electrochemistry was important, but John Dalton's 1803 proposal of the atomic model for the structure of substances was more important; the proposal was supported, extended and partially consolidated by Jön Jacob Berzelius, an extraordinary Swiss chemist who gathered the greatest amount of atomic (and molecular) weights. Berzelius discovered several elements and suggested the electrochemical model of the chemical bond and the nomenclature of the elements that we use today.

In those days, not everyone accepted the existence of the chemical atom as a chemical entity. For instance, in France, the model of equivalents allowed an explanation of chemical reactions (mainly those of neutralization) without one having to accept the atomic model.

In 1828, the synthesis of urea by Friedrich Wöhler and the concept of isomerism heralded the beginning of organic chemistry. Since then, not only has the composition of chemicals been valuable, but the structure has also become very important. Shortly thereafter, Berzelius separated chemistry into inorganic and organic, the latter being interpreted with Charles Frédéric Gerhardt's types model, instead of the electrochemical bond model proposed by Berzelius himself. The electrochemical model indicated that the formation of all substances was due to the attraction between opposite electric charges, and it was successful with inorganic substances but a failure with organic substances. For this reason, many chemists accepted the vitalist model that held that through a vital force, different from the principles of inanimate objects, such as salts and inorganic minerals, complex organic substances were formed. However, this model collapsed with the synthesis of urea. Thus, for many of the following decades, chemical practices were separated by those who worked with inorganic substances and those who did so with organic substances.

The distinction between atoms and molecules was established in Karlsruhe, Germany, in September 1860, at the First International Congress of Chemists. The meeting was convened by some renowned personalities of that time who yearned to reform and enhance the language of chemistry. The purpose was not achieved, but after the meeting, some participants shared different *exemplars* that materialized in textbooks. Two of those exemplars related to analytical instruments: the *kaliapparatus*, developed by Justus von Liebig for the determination of the minimum formulas of the organic substances, and the polarimeter, used by Louis Pasteur to characterize optical isomerism until then only identified in organic substances. Chemistry was a European public activity. Different models were developed to explain new reactions and new chemical properties, like aromaticity. On the other hand, and since the participation of Stanislao Cannizzaro in Karlsruhe, molecules were clearly differentiated from atoms. Dmitri I. Mendeleev, another attendee to the event, using the valence model and atomic weights, built his famous periodic table. Until today, different models developed around the original concept of valence played a significant role in chemical practices.

In 1874, independently of each other, Jacobus Henricus van't Hoff and Joseph Achille Le Bel explained optical isomerism in molecules from the asymmetry of the carbon atom. Since then, the molecule has been the most important chemical entity. Many of the theoretical doubts were dispelled, while the industrial advances due to the discovery of the mauve dye by William Henry Perkin accelerated, particularly in Germany, the creation of a transnational chemical industry. Meanwhile, in England, the Alkali Act of 1863 was published to stop the discharges of hydrochloric acid into the atmosphere. Thus, chemistry, originally inorganic, had a subdiscipline broader than itself, organic chemistry, and this was its first big transformation.

The German university model that closely linked *pure* and *applied* research was copied in many European countries. Among all sciences, chemistry was the first in which experimental work during its teaching became mandatory. At that time, compulsory education was imposed in many European countries, and schools began to be built and managed by governments. Meanwhile, everything that could be synthesized and marketed was being produced. As an example, the respected German pharmaceutical company Bayer did so. In addition to aspirin, they commercialized cocaine and heroin. In Sweden, Alfred Nobel invented dynamite, and controlled explosions began to change the surface of the planet. In 1868, spectroscopy, developed by Robert Bunsen and Gustav Kirchhoff, allowed the discovery of the element helium (He) in the Sun.

Through analysis and synthesis, organic chemists could copy molecules that were originally in plants and animals and then produce entirely new molecules. Substitution models were used to explain specific chemical reactions. European societies first, then the rest of the world, were flooded with new dyes, materials, and medicines from the powerful German chemical industries. In Germany, as the number of colleges and universities with chemistry departments increased significantly, so did the numbers of their teachers, researchers, and students.

Chemists gradually incorporated more instruments in their laboratories and measured more accurately. Thermodynamics were well recognized. Physical chemistry and chemical physics,

the new specialities of both chemistry and physics, was under way. In 1887, Friedrich Wilhelm Ostwald and J. van't Hoff founded the first journal devoted to this subdiscipline. In the United States, chemical engineering emerged around the concept of unit operations. Arthur D. Little recognized that the majority of chemical processes were different combinations of a small number of operations like heating, cooling, distilling, drying, and so on, and his influence grew accordingly.

With his cathode ray tube, the British physicist J. J. Thomson discovered the electron, and the chemical community accepted that atoms could be divided. The research and characterization of isotopes by Francis William Aston using his mass spectrograph, together with the discovery of radioactivity, produced a turning point in the practice of chemistry. The recognition of the existence of atomic nuclei and electrons, the new entities of chemistry, gave rise to models (e.g., by Gilbert N. Lewis and Walther Kossel) that explained the nature of the chemical bond, as well as ions and radicals. In chemical practices, and in its exemplars, a priority of entity realism over theory realism can be recognized: Long before an appropriate theoretical representation was developed, chemists denoted their scientific objects as ions or radicals. The reality of atoms and molecules stopped being discussed. With spectroscopy and X-rays, electromagnetic radiation occupied an important place in chemical thought, increasingly influenced by the advances that were taking place in physics. It was clear that beneath the omnipresent materiality of the chemical substances, there was a reality that only physicists could access by techniques that were being developed.

Against the easy commercialization of any substance, a usual practice at the beginning of the 20th century around the world, in the United States, the Food and Drug Administration was established with the intention of regulating the local food and medicine market, a rule that later became global.

The outcome of the First World War, in which the United States participated decisively, confirmed that world geopolitics were changing. Although Germany lost the war, the German officer Fritz Haber developed the most important

techno-scientific process in the history of mankind, the artificial synthesis of ammonia, a fundamental substance in the production of fertilizers and explosives. Chemical practices are not ethically neutral. This synthesis gathered knowledge from different chemical, physical, and engineering practices. The American atomic model of Lewis faced the European model developed by quantum physicists. The first one could explain the chemical bond, fundamental to chemical practices, whereas the second one, spectroscopy, became fundamental to physics practice. Both models were the result of their time and of the ambitions of their creators. This was the second major transformation of chemistry, which entailed a different feeling about the world.

In 1945 at the end of the Second World War, the president of the National Science Foundation of the United States, Vannevar Bush, published the report *Science, the Endless Frontier*, in which he openly requested the federal government to finance science research in American universities and to give support to the companies that had supplied materials and equipment to the U.S. Army. With the adoption of this proposal, chemical laboratories changed more than they had done in the previous 300 years. From that moment, referred to as the *instrumental revolution*, the use of the next instruments became widespread in chemical practices: electrophoresis and ultraviolet, visible, infrared and nuclear magnetic resonance spectrometers; X-ray crystallography; and particularly since 1956, mass spectrometers. Meanwhile, chromatographs came to occupy a place in chemical laboratories' tables. New equipment industries were created following the military logic of the standardization of parts, which facilitated their consumption. The new subdiscipline of instrumental chemistry appeared. Robert Burns Woodward took advantage of the arrival of new instruments, such as nuclear magnetic resonance, in his research on synthetic chemistry. That allowed him, and some others, the preparation and characterization of more complicated products, many of them of great biological and medical importance. The synthesis of morphine, cholesterol, cortisone, strychnine, penicillin, and chlorophyll shared the onset of tranquilizers (such as chlordiazepoxide and diazepam), as well as contraceptives. The total synthesis of vitamin B₁₂ is a milestone in the

history of chemistry. Plenty of chemists in many countries worked for 11 years to synthesize this complex molecule with nine asymmetrical carbon atoms. Besides, commercial macromolecules changed the material world. The post–World War II period marked the beginning of the plastics era. Different research groups around the world undertook the application of thermodynamics and chemical kinetics to the systematic study of these materials.

The new subdiscipline of molecular biology emerged through the integration of chemistry, biology, and physics to answer a question: How is genetic information transferred from one organism to another? Molecular biology was trying to discover general physicochemical models that govern vital phenomena whereby macromolecules, especially proteins, became the principal focus. Since its origin, molecular biology has depended upon the design, provision, and maintenance of complex and expensive instruments.

At that time, computers were incorporated into chemical research practices and with them the programs that allowed chemical calculations. Many chemists, championed by Linus Pauling and Robert S. Mulliken, started to think about the structure of matter in terms of quantum mechanics. They did so considering some of the principles of quantum mechanics, incorporating broadly empirical information from chemistry practices. Quantum chemistry appeared as a new subdiscipline of chemistry, with new educational challenges. Instrumental chemistry, molecular biology, and quantum chemistry, with their own models and practices, characterize the third big transformation in chemistry.

The words *plastic* and *flexible* became commonplace and identified a socially valuable attitude, although the use of plastic objects led to huge consumption and to a massive global pollution problem. The new subdiscipline of organometallic chemistry, with its interest in the function of catalysts (a concept developed by Berzelius at the beginning of the 19th century), consolidated. At the end of the 20th century, the size and type of chemical objects (substances), the way in which they must be produced, and the time in which they are transformed were specified.

In 1974, Frank S. Rowland and Mario J. Molina published the results of their research on the effect

of chlorofluoroalkanes in the ozone layer. It was not the first time that chemical companies and governments around the world faced difficulties for industry's ability to pollute the environment, but at that moment, the damage and the risk were unequivocally global. In this sense, a couple of years earlier, the use of the insecticide DDT in the United States was banned as a result of the extensive information coming from a new instrument, "the most sensitive, easily portable and inexpensive analytical device in existence," the electron capture detector, invented and refined by J. Lovelock. After a strong struggle with the chlorofluoroalkanes chemical industry, where Molina and Rowland played an active role, the political response to ozone layer depletion was the Montreal Protocol. Signed in the 1980s by more than 200 countries, it was the first universally ratified treaty in UN history. So after the publication of Rachel Carson's book *Silent Spring* in the 1960s, the foundation of the U.S. Environmental Protection Agency in the 1970s and the Montreal Protocol, green chemistry as a new subdiscipline was born.

At that moment, the community learned how to make chemical reactions in less extreme conditions (in terms of pressure, temperature, and solvents) than those used hitherto. Chemical practices approached conditions that allow life and decreased the manufacture of potential pollutants. The construction and handling of large and different molecular aggregates gave rise to supramolecular chemistry. Close to this subdiscipline is another one: *nanochemistry*. This subdiscipline refers to the possibility of using chemical synthesis knowledge to build molecular aggregates of specific size, shape, composition, or surface, and it has multiple applications today in medicine, cosmetics, and materials. The origin of nanochemistry is related to the discovery of *fullerenes*, so called, in 1984 by Robert Curl, Harry Kroto, and Richard W. E. Smalley, and the subsequent synthesis of carbon nanotubes. At the same time, physicists at IBM designed the scanning tunneling microscope and the atomic force microscope instruments that allow *seeing* atoms and manipulating them individually at very low temperatures. In 1999, Ahmed Zewail was awarded with the Nobel Prize in Chemistry "for his studies of the transition states of chemical reactions using femtosecond spectroscopy." Although there are

chemical phenomena that last for billions of years, the most basic processes take place at few femtoseconds (1×10^{-15} s). This was the fourth big transformation.

Perhaps the biggest change in the chemical industry took place in pharmacy. Since the second half of the 19th century, many of the European dye companies were transformed into pharmaceutical houses. At the end of the 20th century, more than half of the world's medical research was conducted in the United States. Some of the most important developments in this short period were the use of recombinant DNA technology and the use of combinatorial chemistry for the design of medicines. Furthermore, three Swiss companies (Ciba, Geigy, and Sandoz) merged into the largest global business fusion in history, creating Novartis.

Concluding Remarks

After the four major transformations characterized, among other things, by the appearance of new, multiple, and important subdisciplines, all of them with their own models, chemistry is in need of redefinition. Consider what was stated about chemistry in the original French Encyclopedia, one of the chief works of the Enlightenment era, edited by Denis Diderot (1751, vol. 3, p. 408):

Chemists are distinct people, still very few, with their own language, their laws, their mysteries, and living almost isolated in the midst of great people hardly curious about their business and expecting nearly nothing from their industry. . . .

Clearly, this is no longer the case. In its relatively short history, the contribution of the chemical industries to widespread degradation of the natural environment cannot be ignored and has indeed become a matter of grave worldwide concern. Finally, in the 21st century, a clear and detailed acknowledgment of the history of women's contributions to chemistry, as well as the other sciences, is long overdue.

José A. Chamizo

See also Materials; Models; Pluralism; Practices; Technoscience

Further Readings

- Aftalion, F. (2001). *A history of the international chemical industry: From the "early days" to 2000*. Chemical Heritage Press.
- Bensaude-Vincent, B., Simon, J. (2010). *Chemistry: The impure science*. Imperial College Press.
- Chang, H. (2012). *Is water H₂O? Evidence, pluralism and realism*. Springer.
- Hall, N. (Ed.). (2000). *The new chemistry*. Cambridge University Press.
- Klein, U. (2003). *Experiments, models, paper tools: Culture of organic chemistry in the nineteenth century*. Stanford University Press.
- Kuhn, T. S. (1970). *The structure of scientific revolutions*. University of Chicago Press.
- Mauskopf, S. H. (Ed). (1993). *Chemical sciences in the modern world*. University of Pennsylvania Press.
- Morris, P. J. T. (2016). *The matter factory: A history of the chemistry laboratory*. Reaktion Books.
- Schummer, J. (2010). The philosophy of chemistry. In F. Allhoff (Ed.), *Philosophies of the sciences*. Blackwell-Wiley, 163–183.
- Suckling, C. J., Suckling, K. E., & Suckling, C. W. (1978). *Chemistry through models: Concepts and applications of modelling in chemical science, technology and industry*. Cambridge University Press.
- Tobin, A., & Ambrosio, C. (Eds.). (2016). *Theory choice in the history of chemical practices*. Springer.

CHURCH-TURING THESIS

The Church-Turing thesis, properly understood, is a thesis about the semantic relationship between a group of linguistic expressions. It is a thesis for which there can be no mathematical proof, but it is generally considered to be beyond reasonable doubt and is almost universally accepted by mathematicians and philosophers of mathematics. The thesis is that a certain many-membered set of precisely defined mathematical expressions, which are coextensive with each other, are coextensive with the nonmathematical expression *effectively computable*, as that expression is used by mathematicians. The fact that there are many members of the set of precisely defined mathematical expressions is of utmost importance. To enable readers to understand what the thesis, as stated,

means, the purpose of this entry is to enable them to appreciate the importance of this latter fact.

This entry first outlines the general notion of a function, which it is crucial to understand if one is to understand the Church-Turing thesis. Then, the term *effectively computable* is introduced, followed by a number of the most important members of the set of precisely defined mathematical expressions, viz. *general recursive*, *λ -definable*, and *Turing computable* (these are the expressions discussed by Alonzo Church and Alan Turing in 1936, the year in which they independently introduced the thesis). The meaning of the thesis is then explained using the notions and terms introduced. The term *Turing computable* is then further explicated by giving an example of how the term is mathematically defined, and the entry finishes with a brief discussion of the significance of the thesis, illustrated by a further example given in terms of Turing computability.

Functions

The set of natural numbers, traditionally named N , is the set containing 0 and all of the positive numbers used for counting (sometimes 0 is excluded, but it is included here). So N contains infinitely many members: 0, 1, 2, 3, 4, and so on. For every member of N , there is another member of N that is 1 greater than it, called its *successor*. For example, N contains the number 3, so it contains its successor, the number 4. And N contains the number 3,450, so it contains its successor, the number 3,451. The *successor function* [here called $s(x)$] is the function that associates every natural number with its successor. So, $s(x)$ associates 0 with 1, 1 with 2, 2 with 3, and in general n with $n + 1$. The *domain* of a function is the set of objects for which the function is defined, so as $s(x)$ is defined for all natural numbers, its domain is N . The *range* of a function is the set of things which it associates members of its domain with, so because every natural number other than 0 is the successor of some other natural number, and because $s(x)$ associates every natural number with its successor, the range of $s(x)$ is the set of all natural numbers excluding 0. The members of a function's domain are called its *arguments*, and the members of its range its *values*. Thus, $s(x)$ has the value 1 for the argument 0, the value 2 for the

argument 1, the value 3 for the argument 2, and in general the value $n + 1$ for the argument n .

$s(x)$ has a single argument place. That is, it takes single objects (i.e., individual numbers) from its domain as arguments. But there is no restriction on how many argument places functions can have. For example, the *addition function* (here called $SUM(x,y)$) has two argument places; it is the function that associates every ordered pair of (not necessarily distinct) natural numbers x and y (written as $\langle x,y \rangle$) with their sum. For example, $SUM(x,y)$ has the value 5 for the arguments $\langle 2,3 \rangle$ (in this case, the argument places are filled by two distinct members of its domain) and the value 72 for the arguments $\langle 36,36 \rangle$ (in this case, both argument places are filled by the same member of its domain). Nevertheless, each function has a fixed number of argument places, called its *arity*. So, $s(x)$ has an arity of 1 and $SUM(x,y)$ an arity of 2. Despite it being possible for a function to have n argument places, for any $n > 0$, it is required (by stipulation) that a function with an arity of n associates each n -tuple of arguments with a single member of its range. (In practice, this is no real restriction at all, however, for the range of a function can itself contain ordered tuples of objects, but this will be of no concern in this entry.)

The two examples of functions given above are functions from natural numbers to natural numbers. Despite the fact that there is no restriction on the types of objects that functions can have in their domains or ranges, so shall be the further examples given in this entry. This is because the Church-Turing thesis only concerns those functions that can be represented as functions of natural numbers.

The Term Effectively Computable

The first of the linguistic expressions that feature in the Church-Turing thesis is the term *effectively computable*. This term is used by mathematicians to pick out a certain set of functions, but it is not a precisely defined mathematical term. It is applied to a function when there is an effective method for obtaining the value for any of its arguments. Such a method consists of a list of instructions that could be followed by a human working with only pen and paper that would, in principle and if

followed without error, provide the correct value for any of its arguments. The method must be completely explicit at each step and require no ingenuity nor information other than that contained within its steps to follow. It must be, in other words, the kind of method that could be implemented by an automated device capable of no independent thought. The method must also contain a finite number of steps, each of which must take a finite length of time to perform. Even with these restrictions, however, there are methods that no actual human being could practically perform. For example, there are methods that require more digits to be written down, in any numeral system, than there are atoms in the universe. And there are methods that, despite being finite, would take any actual human working as fast as they could until the heat death of the universe occurs to complete. But these are practical limitations of space and time, and they are ignored when talking about effective computation. In order for a function to be effectively computable, all that is required is that the method would, given enough space and time, and ignoring any other practical difficulties with its implementation, produce the correct value for any of its arguments.

Any effective method must represent the natural numbers using some numeral system or other (so far in this entry, the familiar Arabic numerals—"0," "1," "2," etc.—have been used). So long as there is an effective method for translating between one system and another (as there is for all known systems), it makes no difference which is used. But for practical and illustrative purposes, certain simple systems are often easier to work with. Later, when discussing Turing machines, the tally system will be used, in which a particular number n is represented by $n + 1$ strokes (so 0 is represented by "I," 1 by "II," 2 by "III," etc.). But it can also be used fruitfully now to give a simple one-step example of an effective method for producing the correct value for any argument n of $s(x)$:

1. To find $s(n)$, write down $n + 2$ strokes.

For $s(3)$, for example, one would write down five strokes, "IIII," which represents the number 4. It is intuitively clear that this method meets the requirements laid down for being an effective

method, and so, this shows that $s(x)$ is effectively computable.

Effectively computable is often said to be an intuitive or informal term, and it is sometimes said to be vague. What is certainly true is that it is not a precisely defined *mathematical* term. This is because it is defined in terms of expressions (e.g., *requiring no ingenuity to follow*) that are not themselves precisely defined mathematical terms. This seems to be what is meant by those who say that the term is intuitive or informal, and so, one should have no qualms about admitting that it is so. But it is less clear whether the term is vague. Syntactically, the term is a predicate and so has what is known as an *extension*, in other words a set of things that it applies to. A term is vague, in the full sense of the word, if it has an indeterminate extension, that is, if there are some things such that it is indeterminate whether the term applies to them or not. For example, the expression *bald person* plausibly has no determinate extension; there are some people for whom it is unclear whether or not they are bald, and so, it is indeterminate whether such people fall into the extension of the term. The term *effectively computable* may or may not be vague in the same way, but whether or not it is depends upon contested issues in the philosophy of language and metaphysics that there is not enough space in this short entry to discuss. So, for the purposes of this entry, the possibility that the term is vague in this sense should be merely noted.

The Terms General Recursive, λ -definable, and Turing Computable (Among Others)

In 1888, the German mathematician Richard Dedekind introduced a way of defining a set of precise functions now known as the *primitive recursive functions*, which was studied more fully by a number of other mathematicians in the 1920s and early 1930s (most notably by the Hungarian mathematician Rózsa Péter). *Recursion* in mathematics, understood intuitively, is a procedure whereby new functions are obtained by feeding the values of functions back into themselves as arguments (possibly multiple times), and the primitive recursive functions are defined by just such a procedure. A particular set of basic

functions that can be shown to be effectively computable in a trivial fashion (as $s(x)$ was previously shown to be so) are laid down. Then, two procedures for defining new functions from these basic functions, *composition* and *primitive recursion*, are introduced. These procedures are finite and require no ingenuity to carry through, so any function defined from the basic functions using them is also effectively computable. So, it follows that any function that is primitive recursive is effectively computable, and so every function in the extension of primitive recursive is in the extension of effectively computable. It does not follow, however, that any function that is effectively computable is primitive recursive, so it does not follow that there are no functions in the extension of effectively computable that are not in the extension of primitive recursive. And indeed, a number of functions that are effectively computable but not primitive recursive were discovered in the late 1920s. The most well known of these is known as the *Ackermann function*, after the German mathematician Wilhelm Ackermann, who discovered it in 1928.

In 1934, the great Austrian mathematician Kurt Gödel gave a series of lectures at Princeton University where, developing work by the French mathematician Jacques Herbrand, he extended the recursive method by adding a new finite procedure, *minimization*, to the two existing procedures. Gödel called the functions that can be defined from the basic functions using the new procedure along with the old the *general recursive* functions and showed that the Ackermann function is general recursive. Thus, Gödel introduced a new precisely defined mathematical predicate that was known to have in its extensions all primitive recursive functions and at least some others that are effectively computable but not primitive recursive. However, Gödel did not show that *general recursive* and effectively computable are coextensive (i.e., that they have exactly the same extension), and Gödel remained unconvinced that this is in fact so until after the publication of a seminal paper in computability by the British mathematician Alan Turing in 1936. Before Turing's paper, however, came another seminal paper by the American mathematician Alonzo Church.

In the early 1930s, Church was working on the foundations of mathematics and, along with

his fellow U.S. American Stephen Kleene, developed a formalism called the λ -calculus, in which it is possible to precisely define a set of functions known as the λ -definable functions. The method is again finite and requires no ingenuity to perform, so it was known that all λ -definable functions are effectively computable. Then, in 1936, Kleene showed that the terms *general recursive* and λ -definable are in fact coextensive, in other words, that every general recursive function is λ -definable, and vice versa. This result was followed in the same year by the publication of Church's seminal paper on computability, where the Church-Turing thesis was first introduced (although it is customary to refer to Church's version of it simply as *Church's thesis*). The methods used to define the λ -definable functions and the general recursive functions are very different, but both seemed to Church to be equally natural ways of cashing out the notion of what it is for a function to be effectively computable. So the fact that the two turn out to be coextensive was taken by Church to be strong evidence that they do in fact capture that intuitive notion.

Later in 1936, Church's conjecture was bolstered when the British mathematician Alan Turing published his seminal work on computability. In his paper, Turing laid down yet another finite method for defining a precise set of functions that involved no ingenuity to perform. He thus introduced another mathematical predicate (now known as *Turing computable*) that had as its extension a precise set of functions. Turing's method involves defining an idealized machine known as a *Turing machine*, which carries out computations by following instruction to perform a finite series of simple actions from an explicitly defined list. A function is said to be Turing computable if and only if it can be computed by such a machine. Turing's method, although strictly defined, matches the notion of an effective method more closely than those used to define *general recursive* and λ -definable, which gave it a more compelling *prima facie* claim to capture that intuitive notion than those previous methods. But, moreover, Turing was able to show that *Turing computable* is also coextensive with λ -definable and so with *general recursive* too. So, at this point, three distinct ways of cashing out the intuitive

notion of effective computation were known, all of which gave rise to predicates that turned out to be coextensive.

Since 1936, a significant number of other ways of cashing out effective computation in strict mathematical terms have been laid down, and each has been shown to define a predicate that is coextensive with the predicates *Turing computable*, *general recursive*, and λ -*definable*. It is this fact that has convinced mathematicians and others of the truth of the Church-Turing thesis. But what precisely is the thesis? In its strongest and most simple form, it is simply the thesis that *effectively computable* is coextensive with *Turing computable* (or *general recursive*, or λ -*definable*, or with any of the other coextensive mathematical predicates that have been defined). But the possible vagueness of *effectively computable*, noted earlier, indicates that there is a weaker version of the thesis in the vicinity. For anyone who believes that the predicate *effectively computable* is in fact vague (i.e., that it has an indeterminate extension), the thesis is not, strictly speaking, that *effectively computable* is coextensive with the precisely defined mathematical predicates. Rather, it is that its extension is satisfactorily close to the extension of those mathematical predicates. The most natural way of cashing out the weaker version of the thesis is to hold that (i.) every function that determinately falls into the extension of *effectively computable* also falls into the extension of the mathematical predicates, and (ii.) every function that determinately does not fall into the extension of *effectively computable* does not fall into the extension of the mathematical predicates. On the weak version of the thesis, the meaning of the mathematical predicates (in one sense of *meaning*) is precisification of the meaning of *effectively computable*, while on the strong version, the meanings are the same.

Examples of How Turing Machines Compute Functions

In this brief entry, it is not possible to give an exhaustive account of each of the mathematical methods for defining functions described in the preceding sections. However, as some familiarity with at least one method gives a greater understanding of the thesis, in this section, Turing machines are described in more detail, and two

simple examples of how they compute functions are given.

A Turing machine is an idealized computing device that is most often visualized as a small box on wheels that moves up and down an infinitely long tape that is divided into squares. Each square is marked with either a single cross or a single stroke, and the box has something inside it (it doesn't matter what) that is capable of reading whatever is on the square it is positioned above, and of performing one of five actions: (1) It can delete whatever is in the square and replace it with a cross; (2) it can delete whatever is in the square and replace it with a stroke; (3) it can move the box one square to the right; (4) it can move the box one square to the left; and (5) it can halt. The thing in the box passes through a series of states, and at each moment what it does is determined by whatever is in the square it is reading and what state it is in. Two kinds of states, an *initial state* and an *end state*, can now be defined. An initial state is a state in which the box is positioned over a particular square S ; in the squares to the left of S , there is nothing but crosses; in S and the squares to its right, there may be strokes that form blocks separated by no more than one cross. So, for example, the tape in an initial state might look like this: “. . . XXXI-IXIIIXIIIXXXXX . . .” but cannot look like this: “. . . XXXIIIXXIIIXIIIXXX. . .” An end state is a state in which the box is positioned over a particular square S ; in the squares to the left of S , there is nothing but crosses; in S and the squares to its right, there may be strokes that form a single block, but no more, followed by crosses. So, for example, the tape in an end state might look like this: “. . . XXXIIIIIXXX. . .” but cannot look like this: “. . . XXXIIIXIIIXXXXX. . .” The tally system introduced earlier is now used such that blocks of strokes are taken to represent the natural numbers. So, for example, the series “. . . XXXIIIXX. . .” printed on a portion of tape represents the number 3, and the series “. . . XXXIXIIIXIIIXXX. . .” represents the n -tuple $\langle 0, 2, 3 \rangle$.

A Turing machine computes the value v of an n -tuple of arguments $\langle a_1 \dots a_n \rangle$ for a function f by having a series of instructions that are such that, if it were started in an initial state that represents $\langle a_1 \dots a_n \rangle$, it would halt in an end state that

represents v . Let us take the functions $s(x)$ and $SUM(x,y)$ once more as examples. Consider the following instructions:

State 1: If there is a stroke in the current square, go right and stay in State 1. If there is a cross in the current square, delete it and replace it with a stroke, and go into State 2.

State 2: If there is a stroke in the current square, go left and stay in State 2.

If there is a cross in the current square, go right and halt.

Given these instructions, a Turing machine started in an initial state on a tape representing 3 (e.g., "...XXXIIIIXXX..."), halts in an end state on a tape representing 4 (e.g., "...XXXIIIIXXX..."). If started in an initial state representing 6 (e.g., "...XXXIIIIIIXXX..."), it halts on a tape representing 7 (e.g., "...XXXIIIIIIIIXXX..."). In general, if started in an initial state representing n , it will halt on a tape representing $n + 1$. So, a Turing machine can compute $s(x)$.

As the functions get more complex, so do the instructions needed that enable Turing machines to compute them. Even the instruction for the (relatively) simple $SUM(x,y)$ is fairly complex. But the following description adequately illustrates how a Turing machine can compute $SUM(x,y)$. It starts in an initial state representing $\langle x,y \rangle$ (i.e., on the leftmost stroke of a block of x strokes that is followed by a single cross and a block of y strokes) and halts in an end state representing the sum of x and y (i.e., on the leftmost stroke of a single block of $x + y$ strokes followed by crosses). So, if started over the leftmost stroke on a tape that looks like this: "...XXXIIIIIIIIXXX..." (representing $\langle 3,2 \rangle$), it halts over the leftmost block on a tape that looks like this: "...XXXIIIIIIIIXXX..." (representing 5). One set of instructions that provides this result tells the thing in the box to move right through the first block of x strokes, past the cross that separates it from the second block of y strokes, until it reaches the rightmost stroke in this latter block. It then tells the thing in the box to delete the rightmost stroke and replace it with a cross. It then tells it to travel left back to the cross it passed before, and to delete it and replace it with a stroke, before continuing to the left to halt

over the leftmost stroke in the block of x strokes. As it is possible to provide these instructions using the five actions given, a Turing machine can compute $SUM(x,y)$.

Significance of the Church-Turing Thesis

The significance of the Church-Turing thesis derives from the fact that it has been shown that there are particular functions of natural numbers that do not fall into the extension of the precisely defined mathematical predicates described above. The Church-Turing thesis (in both its strong and weak form) thus entails that these functions are not effectively computable. Both Church and Turing independently discovered such functions and on the basis of their discoveries independently established that a problem posed by the German mathematician David Hilbert, known as the *Entscheidungsproblem*, is unsolvable. The problem was to show that there is an effective method for deciding whether an arbitrary formula, framed in a system of what is known as *first-order logic*, is provable from the axioms of the system. Church and Turing were able to reduce this problem to the problem of showing that certain functions are computable, and so by showing that particular such functions are not computable, they established that there could be no such method.

Probably the most well-known example of an uncomputable function is the *halting function* (called here " $h(x,y)$ "), which Turing showed to be uncomputable in his 1936 paper. This function, $h(x,y)$, can be described fairly simply in terms of Turing machines. So, given that Turing machines have already been described, this entry concludes with a brief informal proof that $h(x,y)$ is not Turing computable and so by the Church-Turing thesis not effectively computable either.

First, note that there are infinitely many possible Turing machines (consider: there is a Turing machine that computes every function in the series $f(x + 1)$, $f(x + 2)$, $f(x + 3)$...). Next, note that Turing machines can be *enumerated*. This means that each of the infinitely many Turing machines can be associated with a distinct natural number. So, suppose E is an enumeration that does just this. Now, we can describe $h(x,y)$ as the function that

has the value 0 for arguments $\langle x, y \rangle$ if the Turing machine associated with the number x by E and started with input y halts and has the value 1 otherwise.

To see that $b(x, y)$ is not Turing computable, start by assuming that there is in fact a Turing machine, T_1 , that computes $b(x, y)$ and so halts in an end state representing 0 if the Turing machine associated with x by E halts given input y , and in an end state representing 1 otherwise. Now, consider another Turing machine, T_2 , that given a single number as input first copies that number on the tape (so, e.g., started on a tape that looks like this: “. . . XXXIII XXX. . .” it first produces a tape that looks like this “. . . XXXIII XXXIII XXX. . .”) before returning to the leftmost stroke, whereupon it follows exactly the same instructions that T_1 follows, except that it has the following instructions in place of those that tell T_1 to halt (i.e., it has two extra states s and $s + 1$):

State s : If there is a stroke in the current square, go right and go into State $s + 1$.

State $s + 1$: If there is a cross in the current square, go left and go into State s . If there is a stroke in the current square, go left and halt.

So, given input x , T_2 behaves just as T_1 does for input $\langle x, x \rangle$, except that when T_1 halts on a tape representing 0, T_2 instead goes into an infinite right-left loop, and so never halts. Now, consider that T_2 will be associated with a natural number n by E and consider what happens if we use T_1 to compute the function $b(n, n)$. There are only two possible results:

A: T_1 halts in an end state representing 0 given input $\langle n, n \rangle$.

B: T_1 halts in an end state representing 1 given input $\langle n, n \rangle$.

Now, given our assumption that T_1 computes $b(x, y)$, we have:

If T_1 halts in an end state representing 0 given input $\langle n, n \rangle$, then T_2 halts given input n .

If T_1 halts in an end state representing 1 given input $\langle n, n \rangle$, then T_2 does not halt given input n .

But now, from what was said about how T_2 operates, we also have:

If T_2 halts given input n , then T_1 halts in an end state representing 1 given input $\langle n, n \rangle$.

If T_2 does not halt given input n , then T_1 halts in an end state representing 0 given input $\langle n, n \rangle$.

So from A, together with (i.) and (iii.), we can derive:

A*: T_1 halts in an end state representing 1 given input $\langle n, n \rangle$.

And from B, together with (ii.) and (iv.), we can derive:

B*: T_1 halts in an end state representing 0 given input $\langle n, n \rangle$.

So, because A* contradicts A, and B* contradicts B, we can conclude that our assumption that T_1 computes $b(x, y)$ must be false. So, no Turing machine can compute $b(x, y)$, so this function is not Turing computable, and so by the Church-Turing thesis, it is not effectively computable either.

Benjamin L. Curtis

See also Axiomatic Theory; Completeness; Gödel's Incompleteness Theorems; Logic, Formal and Informal; Logical Theory; Mathematics, 20th century

Further Readings

- Black, R. (2000). Proving Church's thesis. *Philosophia Mathematica*, 8(3), 244–258.
- Boolos, G. S., Burgess, J. P., & Jeffrey, R. C. (2007). *Computability and logic*. Cambridge University Press.
- Church, A. (1936). An unsolvable problem of elementary number theory. *American Journal of Mathematics*, 58, 345–363.
- Kleene, S. C. (1952). *Introduction to metamathematics*. North-Holland.
- Turing, A. M. (1936). On computable numbers, with an application to the Entscheidungsproblem. *Proceedings of the London Mathematical Society, series 2*, 42(1936–1937), 230–265.

CLIMATE SCIENCE

Climate science is the branch of scientific inquiry that studies the Earth's climate. Roughly speaking, an area's climate is the long-term average of its prevailing weather conditions. Climate science studies how the Earth's atmosphere, oceans, land surface, ice sheets, and biosphere interact in creating its climate. It is thus a field that draws on and integrates different kinds of scientific expertise. In the 20th century, we became aware of our own influence on the climate, which brought climate science to prominence. In particular, the burning of *fossil fuels* such as coal and petroleum releases gases into the atmosphere that have a warming effect on the climate. This anthropogenic (human-caused) warming will have significant negative effects for people and ecosystems around the world, and continued emissions will worsen these effects. Much of climate science from 1970 into the 2020s has been devoted to understanding this climatic change and helping policymakers decide how to mitigate its harm or adapt our societies to it.

This entry continues with an overview of climate science. It then reviews some theoretical questions in the foundations of climate science, such as the definition of *climate* and *climate change*. Climate change is a major social problem, and the following section sets it up as the background for the coming discussion. There follows a discussion of the methods of climate science and theoretical issues associated with them. Next come two major areas of climate policy: mitigation and adaptation. The entry closes with a discussion of late 20th-century climate skepticism.

Overview

The English word *climate* comes from the Greek *klima*, meaning region, and the roots of climate science lie in describing regional variations in temperature and habitability. In this sense, writing on the variation of regional climates goes back at least as far as ancient Greece and the writings of Herodotus. But the origins of modern climate science are usually dated to the 17th century, with the invention of the modern thermometer and barometer. It is around this time that

climate came to mean roughly what it does today: a pattern of weather.

Modern climate science studies the Earth's climate system. According to the Intergovernmental Panel on Climate Change (IPCC), the climate system is made up of five major components which interact to create the Earth's patterns of weather. The first component is the atmosphere: the layers of gas surrounding the planet. Second, the hydrosphere: the parts of the Earth's surface covered in liquid water, including all the oceans, seas, and lakes. Third, the cryosphere: the parts of the planet covered in ice, including the ice sheets, sea ice, lake ice, and snow cover. Fourth, the lithosphere: the rocky crust of the Earth itself, its tectonic plates, their movement, and associated seismic and volcanic activity. Fifth, the biosphere: the collection of all the Earth's living organisms and their ecosystems. Life requires certain conditions in the atmosphere, hydrosphere, cryosphere, and lithosphere and in turn shapes them.

The climate system changes over time, owing to the interaction between these five systems and because of *external* influences upon it. Climate science focuses on three external influences: solar radiation, volcanic eruptions, and human activity such as greenhouse gas emissions and changes in how we use land. Of these, only the Sun is truly external to the Earth's climate system, but this is the standard usage.

Climate science aims to explain and predict the events, changes, and workings of this climate system. In the second half of the 20th century, climate science developed rapidly and gained in prominence. This was due to the development of computers and the discovery of anthropogenic climate change. Computers, with their capacity for rapidly processing massive amounts of data, made it possible to simulate the dynamics of the atmosphere and ocean, creating new capabilities for explaining and predicting climate phenomena. Anthropogenic climate change is a global challenge that has created an urgent need for understanding of the climate system.

Foundational Questions

What is the definition of *climate*? While the vague notion of *prevailing weather conditions*, or *pattern*

of *weather* has been in use since the 19th century, it is difficult to make this definition precise.

The most common definition, as noted earlier, is that a region's climate is its *average weather conditions* over a suitably long period, typically 30 years. While the weather in Hyderabad at mid-day on March 1, 2021, is characterized by a temperature of 33°C, humidity of 25%, northerly wind of 2 m/s, and so on, Hyderabad's climate is an aggregate of these weather conditions. It might be characterized by describing average temperatures (or perhaps average daily highs and lows), average humidity, rainfall, and so on, for a time of year. These quantities (temperature, precipitation, etc.) are known as *climate variables*. It is standard to calculate these averages over 30 years and to use the recent historical record. So, the climate in Hyderabad in 2021 might be calculated from the historical records for climate variables in Hyderabad from 1990 to 2020.

One theoretical problem with this definition is that it can fail to capture marked climatic changes. The eruption of the volcano Mount Tambora in 1815 was the largest in human history: The ash, sulfuric acid, and water thrown into the atmosphere reflected enough of the Sun's radiation to cool the Earth. The eruption coincided with a natural low point in the Sun's radiation, and the combined effects led to the global average temperature falling by approximately 0.5°C. Over the next 3 to 5 years, crops failed due to the cold or the disrupted rainy seasons, which in turn led to rising food prices and famines. The 1810s are the coldest decade on record because of the 1815 eruption. It is natural to say that these were *changes of climate* and that the change occurred in 1815. But the 30-year average definition can't capture this kind of dramatic shift. At best, it allows for a gradual change, as the years 1815 through 1820 are included in the 30-year average.

This and other problems have led to a proliferation of technical definitions of climate and climate change of different kinds. One kind are *distributional* definitions: Climate is a distribution of climate variables over a period of time, subject to certain constraints concerning the internal and external conditions. Another kind are *ensemble* definitions, which also define the climate as a distribution of climate variables, but rather than

being a distribution over the past 30 years, this is a distribution over possible climates which are consistent with what we observe now. A third approach is common among climate physicists, who define climate relative to *models* of the Earth's climate (discussed below). These definitions reference the conditions in models rather than in real or possible climate systems, for example, by defining a climate as an attractor of a non-linear model.

The different definitions exist because scientists want different things from a definition of climate. Distribution definitions have been criticized for being overly statistical, and not referencing physical processes like solar energy flux and ocean currents, which drive climate processes. Ensemble and model definitions are criticized for being too abstracted from the Earth and its actual climate.

A related question is how to characterize the state of the climate. Scientists debate which quantities should be used, either in the distributions that feature in some definitions of *climate* or as a shorthand for talking about climatic change. For example, much focus is given to global mean surface temperature. Targets for climate change are frequently framed in terms of restricting the increase in mean *global surface temperature* (GST) to below some level, such as 2°C above preindustrial levels. But this single quantity is standing in as a representation of the state of the whole, very complex, climate system. There is a range of potential issues here. The first is that a single quantity is unlikely to serve as a good proxy for the state of such a complex system. Communicating about climate change and climate policy goals in terms of it may therefore be too crude or misleading. The second is that GST may not be the best way to represent even this aspect of the climate system: How much heat is accumulating in it? Instead, we may do better by focusing on heat changes in the oceans. But while this may better track some important physical processes, it worsens a third issue with GST: It is abstract and far removed from the changes of climate that people care about, and which policy is interested in mitigating or adapting to. These tend to be local changes, and in particular, damaging extremes of temperature or precipitation.

Climate Change as a Social Problem

Much of the interest in climate science since the mid-20th century is because of changes in the climate. According to the IPCC's fifth assessment report, released in 2014, changes in climate have caused impacts on natural and human systems on all continents and across the oceans. The most cited effect is an increase in GST: From 1880 to 2020, the Earth's temperature increased by about 1°C. This change may sound small, but it belies wide-ranging and significant changes at regional levels. Changes in extreme weather events from 1950 through 2020 have been attributed to climate change, including high temperatures, high sea levels, and heavy precipitation. These events have hurt or killed many people and damaged property. There have also been changes to precipitation patterns and melting ice. Many species have changed their geographic ranges and seasonal activities due to such changes. Agricultural crop yields have suffered negatively, and while there are regions in which agriculture has benefited from the changed climate, the negatives are more frequent than the positives.

Climate science has established that human activity is causing these climatic changes. Producing and burning fossil fuels emits gases such as carbon dioxide and methane. These are called *greenhouse gases* because they contribute to the greenhouse effect in which the atmosphere traps heat from solar radiation. While the greenhouse effect is natural and required for life on Earth, adding greenhouse gases to the atmosphere amplifies the effect and leads to a warming climate. This presents us with a social problem of global magnitude. Greenhouse gases are by-products of industrial processes that have been regarded as necessary for human life and flourishing since the industrial revolution. For example, most of our energy has been produced by burning fossil fuels. Limiting further climate change will require substantial and sustained reductions in emissions, which in turn will require significant changes in how we produce energy, manufacture goods, and transport ourselves and our goods.

Climate change is a collective action problem. This term describes a situation in which the best outcome is achieved if all individuals (and companies, countries, and communities) cooperate.

However, there are incentives and competing interests between individuals, which mean that this cooperation is difficult to secure. Here is a simple example: Reducing emissions from an industrial process like cement manufacturing is costly in the short term. So, if other firms begin to decarbonize, a cement firm that continues to use the polluting process will become comparatively more profitable. Securing cooperation can therefore require coordinating effort from outside the industry, in the form of government regulation. Because climate change is a global problem, this requires coordination between governments.

However, countries themselves have disagreed significantly over how to decarbonize. This is because of concerns about historical inequalities between countries, in terms of how much they emitted in the past and how this relates to injustices such as colonization. This is discussed in the Mitigation section later in the entry.

Another issue in addressing climate change is the fact that the harms of climate change will occur mostly in the future, while the costs to decarbonize must be paid now. There is a misalignment between which generations caused the problem, which must pay the price to fix it, and which will suffer the consequences. An important ethical question is thus: How do we weigh the interests of future generations against those of people living today? There is also a political problem: Leaders in democracies are accountable to voters who exist and vote today. Even if future generations' interests ought to be considered, it is unclear how to do so within the decision-infrastructure of early 21st-century democracy.

Methods of Climate Science

Observing Climate Change

Identifying climate change and projecting future changes to the climate require that we understand what it means for the climate to change. *Climate* is an aggregate notion, and the climate system is naturally varying all the time, so defining and detecting a change in climate is not straightforward. The IPCC defines *detection of climate change* as the process of demonstrating that climate has changed *and* that this change is unlikely to be due to internal variability. For example, one

might say that a change is detected in observations if the likelihood of that change occurring by chance due to internal variability alone is determined to be below 5%.

Detecting a change in climate thus requires something like the following process. First, one needs to estimate the internal variability of the climate system (to exclude this as a cause of any observed changes). This is made difficult by most of our reliable data coming from the 20th century, a period in which there were significant external changes, such as increasing emissions. As a result, estimates of climate variability are often obtained from models: The internal variability of the climate in a model is used as a proxy for the internal variability of the Earth's climate. But global climate models often show different levels of internal variability, which makes it difficult to settle on a standard for detection.

Nevertheless, as the effects of climate change grow, it becomes easier to detect. In 2020, a research paper published in *Nature Climate Change* made a startling announcement in its title: "Climate change now detectable from any single day of weather at global scale."

So what data are used to observe climate change? First there are meteorological observations. These are local weather data from the meteorological stations around the world. For example, temperature measurements are aggregated to calculate the GST, which is how we know that the GST has increased by 1°C from 1880 to 2020. Another source of direct measurement is the extent of glaciers. Globally and cumulatively, the thickness change of glaciers has decreased by about 12 meters between 1960 and 2010. The shrinking extent of Arctic sea ice, usually measured at its September minimum and March maximum, is another source of direct observation of global warming.

These data were all gathered relatively recently and so can only give us evidence of climate change in the period after around 1880. Climate scientists want to understand the state of the climate at earlier times and therefore use various proxy sources of data to make inferences about the historical state of the climate. For example, climate scientists study *ice cores*—long vertical pillars of ice cut out of ice sheets. These polar ice sheets formed from thousands of years of accumulated snowfall.

The weight of successive snowfalls compressed snow into ice, forming layers of ice that act like tree rings. By drilling deep (up to 1.5 kilometers), researchers can study how the composition of the ice changed and use this as a proxy for how the climate changed. Air trapped in the ice provides tiny samples of the atmosphere at the point when the ice layer formed. Changes in the presence of certain isotopes of oxygen in different layers gives an indication of changes in air temperature. Particles in the ice provide evidence of the composition of the atmosphere and past events like volcanic eruptions.

There are many theoretical challenges with such proxy investigations of the paleoclimate. One, which concerns ice cores, is limited geographic coverage: Ice sheets exist only at the poles. So, an inference is needed to go from proxy evidence about the past climate at the poles to the past climate of the planet. Another issue is that proxies often reflect the influence of many climate variables at one. Tree rings reflect all the influences on the tree's growth: the soil, precipitation, temperature, sunlight hours, and so on. This makes it difficult to use these proxies to study single climate variables and to calibrate climate models. Third, combining information from different proxies is challenging: Tree rings and ice cores give evidence about different regions, in different periods, and with different temporal resolution. It is a significant challenge to work out how to combine these data and what the resulting uncertainties are.

Models of the Climate

One of the most important tools of climate science is the simulation model. A model is a representation of a system, in this case the Earth's climate system. Climate science uses a spectrum of models, with different degrees of complexity and resolution.

At the most complex end of the scale are *general circulation models* (GCMs), which model the circulation of the atmosphere and oceans. They simulate this circulation in three dimensions and include interactions with the land surface and cryosphere. This is a difficult and complex process, which requires multiple forms of expertise and significant computing resources.

We can think of these complex simulation models as having the following parts. First is a structural–dynamical description of the system: a description of how it works and how it evolves from one state to another. For GCMs, this is a representation of the circulation of the atmosphere and oceans using equations from fluid dynamics. But, as noted earlier in the entry, the climate system also involves the land surface, cryosphere, biosphere, and the remainder of the hydrosphere, each of which has important dynamics. One major trend in climate science in the 2000s has been the development of more complex models capturing more of these factors.

These dynamical relations are captured with differential equations, involving variables (the main quantities of interest, like temperature and precipitation) as well as parameters (which are fixed quantities, supplied exogenously). Models are simplified and idealized representations and so do not attempt to accurately capture every aspect of the climate system. Instead, they use bits of theory—typically physics, in the case of climate science—to capture the most important dynamics of the climate.

The resulting equations involve variables that are continuous functions of space and time and are very difficult, if not impossible, to solve exactly. To solve them, scientists *discretize* space and time and create an algorithm for approximating the solution to the continuous equation. Spatially, this involves dividing the surface of the Earth into a grid. Temporally, it involves thinking of time as coming in discrete *steps*. One specifies the state of the system at a point in time and then calculates how the system will be after one step, at the next point in time. To do this, the continuous *differential* equations describing the atmosphere and oceans are transformed into discrete *difference* equations. Computers, using step-by-step methods, can then produce approximate solutions. The computer programs that explore this approximate behavior of the system are called *computer simulation models*.

This structural–dynamical part of the model captures parts of the system’s dynamics that the modeler (a) has a theoretical understanding of and (b) regards as important to their current purposes. Other parts of the system might be *parameterized*—that is, represented in the model not via a full

dynamical description, but in a simplified form. These simplified representations often take the form of a single number or set of numbers whose values are supplied by measurements, simple calculations, or other models. Clouds are an important example of a feature of the climate which has been treated via parameterization in GCMs.

Finally, a simulation requires a set of initial conditions, specifying the state of the system at the time the model is initiated. The structural–dynamical description, parameters, and initial conditions are sufficient information to *run* the model, that is, to perform the calculations required to specify later states of the system.

Simulation models raise interesting theoretical issues. First, the study of simulation models is a kind of indirect inquiry: Scientists study the models to learn about something else, the climate. Work is required to establish that this is a reliable form of inquiry. Scientists try to establish that their model does accurately represent the target system, by reproducing some of its known behavior. Typically, this involves simulating the past and seeing whether the model correctly *predicts* what happened. But if the climate itself is changing, does successfully simulating the past give us a reason to believe that the model’s predictions of a changed future climate will be accurate? The hope is that the model captures the mechanism driving the climate system, so that when the model is supplied with the right boundary conditions—in this case, the current and expected emissions—it will correctly simulate the future states of the climate. To test this, scientists use paleoclimate data (such as ice cores) to see whether the models capture the mechanisms that underlay past changes in the climate.

Second, it is common to encounter a range of different models representing the same system.

Such models might disagree over the structural relations in the system or over the values of parameters or initial conditions. A prominent example is the sixth Coupled Model Intercomparison Project (CMIP6) ensemble of global climate models. In some cases, such model *ensembles* may indicate disagreements among scientists; in others, they reflect variation in acceptable methods of model construction. In either case, the ensemble represents some of the scientific uncertainty about the climate system. Ensembles pose a

challenge for users of models, such as policymakers, who need to decide which model result to use for their decision.

One approach is to use many models at once, as in CMIP6. Such studies compare the results of different models and attempt to use the distribution of model results to generate more reliable results. There are different approaches to doing this. One is to use the average result. This could be computed by weighing all model results equally and calculating a simple average. Or the models could be weighed differently, perhaps according to their performance in predicting past data or based on their perceived reliability. In either case, there is debate as to why the *average* should be used. Another option is to look at the range of values spanned by the models and to act as though any of these is equally possible. This in effect turns the problem of multiple models from an epistemic problem (Which output is the right answer?) to a practical question (Which answer should one use for deciding?).

Climate Policy

Mitigation

One kind of policy targeting climate change is called *mitigation*. This refers to reducing future changes in the climate by targeting the causes of climate change. The goal, as the United Nations Framework Convention on Climate Change puts it in Article 2, is to stabilize “greenhouse gas concentrations in the atmosphere at a level that would prevent dangerous anthropogenic interference with the climate system.” In practice, this means reducing the release of greenhouse gases into the atmosphere, either by reducing sources of these gases or by increasing the capture of these gases. According to the IPCC, about 78% of the total greenhouse gas emission increase from 1970 to 2010 was due to carbon dioxide from the combustion of fossil fuels and industrial processes. Major sinks for these gases include the forests, soil, and oceans.

To reduce our greenhouse gas emissions, we will need to find substitutes for the burning of fossil fuels. In energy generation, this requires replacing coal- and gas-fired power stations with nonemitting sources of power such as solar,

wind, and geothermal energy. Reducing energy use is also important, as it will allow us to reduce emissions while we are still dependent on fossil fuels. Energy-efficient buildings, vehicles, and industrial processes are thus important areas of mitigation policy.

The development of *carbon sinks* is another important area. A carbon sink is any system that absorbs greenhouse gases from the atmosphere or prevents additional emissions. Often, discussions about developing carbon sinks also include avoiding processes that remove natural sinks, such as deforestation, desertification, and the loss of certain oceanic ecosystems. But this simply avoids worsening the problem rather than actively improving it. It is also important to increase the amount of carbon taken out of the atmosphere to develop new and expand existing carbon sinks. One option is carbon capture and storage, which refers to capturing the carbon dioxide released from large sources, such as power plants and industrial facilities, and storing it rather than allowing it to be released into the atmosphere. Another option is removing carbon dioxide that is already in the atmosphere through technological means.

Who should reduce their emissions, and who should pay for the development of carbon sinks? This is a question that goes beyond climate science and into the realm of climate justice, as we are asked to consider the just distribution of burdens in responding to this global challenge. Countries’ past contributions to the accumulation of greenhouse gases in the atmosphere are different and reflect their histories of industrialization and colonization. The G8 nations (the United States, Canada, United Kingdom, Germany, France, Italy, Russia, and Japan) are regularly attributed with more than 60% of all emissions to 2015, and some studies attribute as much as 85% of all accumulated emissions to them. Residents of these countries have some of the best living standards in the world, a fact that is deeply connected to their harmful emissions. By contrast, the least developed countries on Earth have not emitted very much, and many have argued that if they were to refrain from emitting activities, this would hinder their ability to develop and thereby ensure that their residents would never achieve comparable quality of life.

However, in the early 21st century, the pattern of emissions has shifted. India and China were in the top five CO₂ emitting countries in each year from 1991 to 2020, and China overtook the United States as the world's largest emitter somewhere around 2005. If one accounts for deforestation as a form of emission, Brazil and Indonesia have been top 10 emitters in the 2010s.

Different regions will face different kinds and levels of climate change, some of which will be more harmful than others. Historically marginalized people are expected to experience some of the worst consequences. Current climate projections expect developing countries in the Southern Hemisphere to receive some of the worst effects of climate change. The changing climate will also push people into poverty, as the climatic conditions supporting their livelihoods change. At an individual level, poorer people have fewer resources to recover from hardships and often lack insurance that could assist them. At a group level, historically disadvantaged populations are often last to receive relief after catastrophic events like hurricanes and floods. So, both internationally and intranationally, those who are least responsible for climate change are expected to face its worst consequences.

Adaptation

The second major category of policies for dealing with climate change is adaptation to the expected future climate. The goal of adaptation is to reduce our vulnerability to changes in the climate that might prove harmful, such as changes in the frequency and severity of extreme weather events, or sea-level rise. Adaptation efforts also aim to benefit from any opportunities afforded by climate change, such as changes in precipitation and a longer growing season. As the climate will change in very different ways in different places, adaptation policies are highly varied and location specific.

Adaptation efforts include changes to ecological, social, and economic systems as well as the development of infrastructure. In an area that is expected to experience significantly greater rainfall, for example, adaptation may involve improving stormwater drainage in the cities, altering

riverbanks and building flood defenses. For an area that is expected to experience tropical cyclones for the first time, adaptation may involve the development of early warning systems, changes to building codes, and the establishment of evacuation procedures. Adaptation is a necessary part of any climate policy: Even if emissions were to suddenly stop, the effects from already accumulated gases would continue to be felt for thousands of years.

Contestation and Skepticism

Climate skepticism refers to doubt in the social problem of anthropogenic climate change. It can take one of three forms. The first is doubt in the trend of global warming. The second is doubt about the attribution of this trend to human causes. The third is doubt about the impacts of climate change. Skeptical claims about the trend of global warming include the claim that the measurements are faulty because of the urban heat island effect. Skeptical claims about attribution include the claim that the warming observed is a product of natural variation and the claim that carbon dioxide does not lead to warming. Skeptical claims about impacts include those who think that warming will be entirely beneficial or that humans will naturally adapt to whatever changes occur without much disruption.

Particularly in the United States, the issue is politicized. Conservative and libertarian think tanks and lobbyists have been significant promoters of climate skepticism. In the 2010s, Republican Party voters were less likely than Democrats to rate climate change as a priority issue, and Republican politicians frequently questioned the reality of climate change. This may be the result of prioritizing ideological commitments to small government and the free market over the factual question of whether anthropogenic climate change is happening. Climate change is a global issue, and the kinds of mitigation and adaptation policies mentioned above have been associated with significant state intervention in the form of regulations, taxes, subsidies for new technologies, and large infrastructure projects. (These need not be the only tools for accomplishing mitigation or adaptation, but they

have been significant in proposals in the period 2000–2020.) Conservative opposition to climate action thus seems to follow from a sort of perverse syllogism: If anthropogenic climate change is real, then significant government action is necessary to enact the required mitigation and adaptation policies. But government action is never necessary! So anthropogenic climate change is not real.

Although these are political issues, they are of significant concern to climate scientists, science advisors and communicators, and academics studying the science–policy interface. Climate change is one of the great challenges of our time. It is an issue in which scientists have played, and continue to play, a crucial role in identifying the nature of the problem and its potential solutions. And yet a significant proportion of people do not trust these scientists. Political leaders have actively discredited the scientists and their work, thereby undermining efforts to prevent catastrophic climate impacts. We need better ways of negotiating contested political issues, like the role and extent of government, without undermining climate science and climate scientists.

Joe Roussos

See also Complex Systems; Data Models; Geophysics; Measurement; Meteorology; Modeling

Further Readings

- Barry, R. G. (2013). A brief history of the terms climate and climatology. *International Journal of Climatology*, 33(5), 1317–1320. <https://doi.org/10.1002/joc.3504>
- Broome, J. (2012). *Climate matters: Ethics in a warming world* (Norton Global Ethics Series). W.W. Norton.
- Frigg, R., Thompson, E., & Werndl, C. (2015). Philosophy of climate science part I: Observing climate change. *Philosophy Compass*, 10(12), 953–964. <https://doi.org/10.1111/phc3.12294>
- Frigg, R., Thompson, E., & Werndl, C. (2015). Philosophy of climate science Part II: Modelling climate change. *Philosophy Compass*, 10(12), 965–977. <https://doi.org/10.1111/phc3.12297>
- Intergovernmental Panel on Climate Change. (2014). *Climate change 2014: Synthesis report. Contribution of working groups I, II and III to the fifth assessment report of the intergovernmental panel on climate change* (Core Writing Team, R. K. Pachauri, & L. A. Meyer, Eds.). IPCC. 151 pp.
- Intergovernmental Panel on Climate Change. (2013). *Climate change 2013: The physical science basis. Contribution of working group I to the fifth assessment report of the intergovernmental panel on climate change* (T. F. Stocker, D. Qin, G.-K. Plattner, M. Tignor, S. K. Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex, & P.M. Midgley, Eds.). Cambridge University Press.
- Katzav, J., & Parker, W. S. (2018). Issues in the theoretical foundations of climate science. *Studies in History and Philosophy of Science. Part B*, 63, 141–149. <https://doi.org/10.1016/j.shpsb.2018.02.001>
- Lloyd, E. A., & Eric Winsberg, E. (Eds.). (2018). *Climate modelling: Philosophical and conceptual issues*. Palgrave Macmillan.
- Oreskes, N., & Conway, E. M. (2010). *Merchants of doubt*. Bloomsbury Press.
- Sippel, S., Meinshausen, N., Fischer, E. M., Székely, E., & Knutti, R. (2020). Climate change now detectable from any single day of weather at global scale. *Nature Climate Change* 10, 35–41. <https://doi.org/10.1038/s41558-019-0666-7>
- Werndl, C. (2016, June 1). On defining climate and climate change. *The British Journal for the Philosophy of Science*, 67(2), 337–364. <http://dx.doi.org/10.1093/bjps/axu048>
- Winsberg, E. (2018). *Philosophy and climate science*. Cambridge University Press.

COGNITIVE SCIENCE

Cognitive science is an interdisciplinary domain of investigation that explores how people perceive, understand, and store information about reality. It includes the study of psychology, anthropology, sociology, information technology, biology, linguistics, and neuroscience. It both benefits from these disciplines and contributes to their development. The interdisciplinarity and interrelatedness have led to a plethora of approaches as far as analyzing cognitive science is concerned. The most important of these are discussed in the sections that follow.

Cognitive Science and Methodologies

Systemic Approaches

Cognitive studies can be characterized by a diversity of approaches and tools that are used to conduct research in this discipline. First, since cognition itself involves the work and engagement of different systems, cognitive science can be described as the ability of various systems to cooperate. Systems in this case can be understood in different ways. One of them is to look at systems as a set of different disciplines that, together with their methodologies, form what is commonly known as *cognitive science*. The second way of looking at this issue is to investigate systems as sets of different individual and social identities, concentrating on micro-systems (e.g., one person) and/or more macro-systems (e.g., communities, groups). The systemic perspectives offer the opportunity to study diversified factors of different origin and their roles in creating, sustaining, or eliminating a given entity. Systemic theories furnish the possibility of studying cognitive science as an open system, depending on the environment and responding to its alterations; an open system utilizes inputs from the surroundings and sends outputs to the encompassing environs. Thus, systems can be investigated by taking their boundaries into account. As far as cognitive science is concerned, a demarcation line is difficult to establish due to the interdisciplinary character of methodology and scope of interest. However, boundaries can be observed between subdisciplines, each having some outstanding features and characteristic scientific approaches. The next option is to study systems through their processes. Thus, the dynamism of cognitive science can be investigated by observing various actions and events in systems.

Network and Hybrid Approaches

Apart from systemic approaches, network perspectives also focus on showing the complexities and interrelations of elements. An example is *actor-network theory* (ANT), one among several popular network perspectives, which stresses the role of living and nonliving entities in determining the performance of people and things. As previously stated, ANT can be used to study the

complexities of modern systemic realities. Narrowing the scope of interest to cognitive science, ANT can be used to discuss the role of living (e.g., workers, stakeholders) and nonliving elements (computers, telephones, etc.) in creating and sustaining cognition. *Social network analysis*, another network theory, stresses the relations between people taking part in any cognitive activity, paying attention to such processual issues as knowledge flow, communication, and innovation. It reflects the popular postmodern approach to the concept of cognition, with such aspects as changeability, interconnectedness, and relations being of crucial importance. By applying this approach, one may observe how social networks are created by different factors at individual and group levels and how they influence individual and group cognition.

Hybridity offers another perspective to cognitive studies. The aim of this approach is to highlight how each of the aforementioned stages determines the existence of the other one as well as their mutual performance. Cognitive studies as a discipline is shown not only as the sum of other layers but rather as a hybrid structure, with independent and unique features created by the coexistence of diversified elements.

Fuzzy-Trace Theory

One of the methodologies that can be used in examining cognitive behavior is the fuzzy-trace theory. As discussed by Valerie Reyna, meaningful parts are encoded in memory in two ways: verbatim representation and gist representation. Verbatim representation is related to an objective relation of what has taken place, whereas gist representation is a subjective perception of what happened. Gist representation depends on one's identity, gender, social role, life experience, and genetic aspects; the same information can be perceived differently by two people regarding their attitude toward risks.

Heuristics and Schemas

Heuristics, on the other hand, enables one to find the needed answers to important questions or the best methods to deal with novel situations. Cognitivists are also interested in studying framing, which concerns whether a given piece of

information is presented in a positive or negative way. In that case, one's cognition is guided by the way risks or gains are visualized.

Schemas theory constitutes another key issue in cognitive science; schemas (schemata) encompass a set of concepts that facilitate the gathering, understanding, storing, and using of knowledge by arranging information in meaningful groups. Schemas can be subdivided by taking into account their function, such as describing people (person schemas), social roles (role schemas), or situations (event schemas).

Purpose

Cognitive studies could also be researched by taking into account the focus of investigation. For example, cognitivists study deficits in cognitive abilities. Such investigations facilitate the understanding of various illnesses such as Alzheimer disease and help patients who suffer from cognitive deficiencies. Apart from the aforementioned medical scope of investigation, cognitive science is used for commercial purposes. For example, such areas of study as advertising or branding benefit from knowledge of how customers perceive and understand marketing communication and make it possible to prepare, for example, commercials or advertisements that will meet the needs and expectations of the target audience. Another way of classifying cognition is by taking into account the type of situations; for example, risky situations involve different types of cognitive behavior. One of the key notions is memory that determines the way a concept is perceived and stored. Another field of investigation is motivation. This issue, being pertinent to such disciplines as education and management, facilitates the understanding of how to make learners or workers more eager to learn novel issues or work more effectively.

Research Scope

Another notion in cognitive studies is the scope of investigation. Cognition can also be researched by analyzing individual or social notions that determine the way cognition takes place. Analyzing the personal dimension, researchers are interested in how an individual perceives and understands reality. At this level, studies concentrate on such notions as gender, age, profession, and social and cultural

background. As far as the macro-approach is concerned, researchers are interested in social cognition, in the way a group perceives and reacts to a given stimulus. Moreover, researchers focus on studying how emotions determine cognition; for example, how people who are happy or stressed perceive a given notion. It should also be highlighted that although human beings constitute the primary focus of research in cognitivism, some researchers investigate the cognitive abilities of animals. For example, the LUCS Cognitive Zoology Group at Lund University in Sweden is interested in *cognitive zoology*, observing the cognitive functions of animals and how animals perform in different environments as well as how they learn and adapt to new conditions.

Linguistic Perspectives on Cognitive Science

Since language is among the most conspicuous representations of cognitive functions, the scope of investigation in linguistics is mainly on the cognitive rather than on the physical side of acquiring, understanding, and using language, being determined by individual and social factors. The long-standing interest in cognitivism (the latter being briefly defined as the science of studying the mind) has given rise to an area of studies called *cognitive linguistics*, focusing on mind-related processes connected with coding and decoding language-related content. The discipline of cognitive linguistics is studied in this contribution in two ways. First, the disciplinary perspective is applied, with such subfields within cognitive linguistics to, among others, language learning and teaching, translation, psycholinguistics, neurolinguistics, and sociolinguistics.

Language Education

Language learning and teaching focuses on how children learn their mother tongue and how one's native language accompanies the growth and maturation processes of individuals. Thus, it can be stated that cognitive linguistics has a developmental side, concentrating on how language changes together with an individual. This approach allows studying how people use language in different life situations and how various factors

determine the way one's linguistic repertoire is used. It is represented in, for example, the way individuals learn corporate discourse and professional jargon as well as in the introduction and execution of corporate linguistic policy in companies and organizations employing multilingual staff. The educational aspect of cognitive linguistics is also visible in the focus on teaching and learning foreign languages. The cognitive perspective makes it possible to understand how novel wording is mastered. Cognitivists interested in foreign language and teaching study the interrelation between a mother tongue and a new foreign language to observe how one's first language determines the acquisition of new tongues. Additionally, such factors as the age, gender, profession, and motivation of learners are taken into account in investigating their cognitive abilities. Cognitive linguistics is also interested in how the selected linguistic tools facilitate or hinder cognition. An example of such a communicative tool is metaphor; researchers investigate how symbolic language determines the understanding of texts. For example, such studies include the research on the role of metaphors in creating and disseminating novel or difficult information to a diverse audience. Another field of investigation is the role of figurative language in cross-cultural communication, aimed at serving diverse language users by relying on commonly known and understood symbols. It should be mentioned that cognitivists may also be interested in nonverbal symbolism, analyzing the influence of pictorial metaphors on the target audience.

Translation

Another subfield of linguistic studies researched by cognitivists is translation; the scope of investigation focuses on, for instance, how equivalents are selected by translators as well as what cognitive processes take place during interpreting.

Psycholinguistics, Neurolinguistics, and Sociolinguistics

Yet another discipline is psycholinguistics, which focuses on how a human being is able to produce, understand, and store a piece of a text or an utterance. A related field is neurolinguistics,

with its focus on how the brain and the nervous system influence the production, perception, and usage of linguistic elements. Some traces of cognitive investigation may be observed in sociolinguistics, which aims at showing the relation between the society and discourse.

Aspects of Language

Apart from the discussed domains, the division of cognitive linguistics into subfields may follow the well-known subcategorization used for linguistics, which stresses different aspects of language: semantics, syntax, and pragmatics. *Cognitive semantics* encompasses the interest of researchers in the way words and phrases are created, understood, used, and stored. Apart from the semantic dimension of cognitive investigations, researchers also study syntax and the way grammar is perceived and understood. *Cognitive syntax* may also prove useful for the studies on foreign grammar acquisition as well as the studies on how children formulate sentences. The last element, *cognitive pragmatics*, studies the relation between meaning and context by applying the cognitive perspective. Cognitivists are interested in investigating how the environment determines the way information is perceived and understood. For example, the same data may be comprehended differently in neutral and risky situations. The aspectual approach offers deeper ways of investigating such subdomains of linguistics as translation or language learning and teaching. Taking the issue of translation into account, semantics provides an insight into such translation elements as terms, whereas pragmatics offers information on the role of context in source and target texts.

Economic Perspectives on Cognitive Studies

Many modern companies are interested in cognitive and behavioral processes taking place at the internal and external level of business entities. The link between economics and cognitive studies is analyzed through such subdomains as cognitive management, cognitive cultural economics, behavioral economics, and cognitive marketing.

Cognitive Management

One field of cognitive interest that has been gaining popularity is the theory of *cognitive management*, pioneered by James Eicher, that focuses on behavioral and cognitive styles of managers to facilitate effective leadership and communication in groups; this area of studies is devoted to finding the best methods of managing people.

Cognitive Cultural Economics

The next domain is *cognitive cultural economics*. In the business world of the 21st century, the business-related disciplines of cognitive studies benefit from cooperation with other areas of research that are cognitive-oriented. One example of such a scientific alliance is the relation between cognitive management and cognitive linguistics. It is visible in the focus on how company discourse is used, perceived, and stored at both the individual and societal levels in companies. Thus, the main sphere of interest is to investigate how different stakeholders perceive the company both as individual human beings and a professional (societal) group. To discuss both the inner and outer dimension of organizational communication, the term *company identity* can be employed to investigate the cognitive complexities of managing discourse. Taking into account the multitude of factors shaping company identity, the focus should be on culture (especially its different elements observable in both individual and group levels) as the crucial identity marker. Regardless of the approach used to analyze the individual or social personality, identity is related to symbols (common values, things, and heroes), language (jargon, slogans, and humor), practice (common rituals, ceremonies, and activities), stories (common myths and legends), and various forms of physical expression (movement and architecture). All the abovementioned elements are created by individuals coming from different cultures and, at the same time, the aforementioned elements create the identity of cross-cultural companies and their representatives. Consequently, managerial cognitivists may analyze in greater detail the communicative aspect of company identity through the prism of cross-cultural differences and show the influence of the abovementioned elements on creating the linguistic dimension of identity in professional

settings. Cognition in that case is studied in a very broad sense, analyzing the response and perception of the entire spectrum of stakeholders, which including, among others, employers, employees, mass media, and residents of the town the company operates in.

Cognitive Economics

Another field of investigation benefiting from the interrelation between cognitive studies and the discipline itself is *cognitive economics*, which focuses on cognition in economic issues and decisions. As in the case of cognitive business, cognitive economics may also rely on the achievements of cognitive linguistics. Given that language shapes both individual and social performance—since it is not only a tool creating the economic reality but also the *recipient* of economic conditions—its dual nature makes it a key factor in any discussion on economy-related topics. For example, one of the economic issues researched in the literature is economic inequality, which can be discussed by taking into account different dimensions. For example, the micro-dimension focuses on the individual level of economic gaps between the person and their economic environment. The meso-dimension, depending on the context, can be related to an organization or a community and its economic position within a broader reality. By contrast, the macro-dimension focuses on the studied country and its place in the global economic world. However, no matter which field of investigation is taken under closer scrutiny, it is language (among other factors) that shapes the performance of individual human beings, communities, groups, organizations, and countries as well as determines their position in the global market. Since language is not only a tool of communication but also an instrument of exercising power or benefiting from some resources, its role in the field of economic inequality and international trade should become the focus of modern investigations as well.

Behavioral Economics and Cognitive Marketing

The elements of cognitive studies can also be traced in *behavioral economics*, which focuses on how different factors related to cognition determine economic decisions and market situations.

Taking the customer and product perspectives of business studies into account, such subfields as *cognitive marketing* can be enumerated. The classification can be taken further, by naming the subcategories of cognitive marketing, such as *cognitive branding* or *cognitive advertising*. The subdisciplines of cognitive marketing encompass the investigation of different cognition types. For example, since branding focuses on such dimensions as the verbal, visual, audio, and olfactory ones, cognitive branding is supposed to research cognition by paying attention to different kinds of stimulus and the overall effect, respectively.

Psychological and Sociological Perspectives of Cognitive Science

The interrelation between psychology, sociology, and cognitive science is visible in at least five subdomains. The first of them, *cognitive psychology*, can be defined as the field of study concentrating on mental processes since it provides an answer to how people perform different functions related to cognition, such as thinking, remembering, and selecting. Another subdomain, *cognitive sociology*, studies how social determinants influence the processes of understanding and perceiving the reality.

The psychological dimension of cognitive studies is also realized by paying attention to the process of decision making: *cognitive heuristics*. In addition, it should be mentioned that the psychological dimension of cognitive science can be analyzed by taking into account both healthy and unhealthy subjects, offering methods of releasing stress or fighting with diseases. Relying on modern achievements in cognitive behavioral therapy and hypnosis, Trevor Silvester introduced *cognitive hypnotherapy* for people whose professions demand high concentration, quick reactions to changing environments, and the necessity to face different stressors. Another field of interest in psychology is the relation between addictions and cognition. In that case, researchers representing mainly psychology and neuroscience investigate how the use of addictive substances determines the cognitive abilities of individuals.

Artistic and Philosophical Perspectives of Cognitive Science

The artistic sphere of cognitive science encompasses the approach connected with different domains of art and the way these types of arts can be perceived and understood. For example, *cognitive poetics* focuses on the application of cognitive psychology to the study of literary texts. *Cognitive aesthetics* concentrates on the emotions and feelings connected with artistic experience. It studies how individuals react to works of art, such as paintings, music, and imaginative literature. This aspect is also crucial for medical applications of art in facilitating clinical treatment or in marketing efforts to stimulate purchasing behaviors. Cognition has interested philosophers from the time of classical antiquity. The current representatives of philosophy who aim to observe the link between cognitive science and philosophy concentrate on the domains referred to as *philosophy of mind* and *cognitive philosophy*.

Intercultural Perspectives of Cognitive Science

Cognition also takes into account differences between cultures. Owing to the advancements related to modern technology and the consequently high mobility of people, ideas, and things, contacts with other cultures and access to culturally different products have become relatively easy. Consequently, a modern human being typically has a hybrid identity composed of different cultural elements and is constantly exposed to contradictory values and points of view. This may sometimes lead to cognitive dissonance, with novel attitudes held alongside, and in contradiction to, the older ones. When faced with a new situation, an individual may focus on cognitive egocentrism, stressing the superiority of their way of thinking over the ones shared by others. Intercultural differences are also connected with different viewpoints on such issues as humor, the role of family in one's life, attitudes to material world, and so forth. In addition, cross-cultural studies provide information on why people perceive and understand various

notions in different ways. For example, such issues as the openness to novelty or the role of tradition can be taken as a marker determining the way various notions are perceived. Cognitive studies in intercultural communication also encompass investigations of differences in perception and categorization of color and shapes. Another notion studied through the prism of cognition is symbolism; there are symbols that are culturally specific as well as the ones that are universal. As far as the methods of studying cognition are concerned, both qualitative and quantitative approaches may prove useful.

Neuroscientific Perspectives on Cognitive Science

Neuroscience constitutes one of the disciplines to which cognitive science is strongly related. The neuroscientific dimension of cognitivism focuses on how the brain and the nervous system determine cognitive and behavioral bodily functions. Applying both micro- and macro-approaches, neuroscientists are interested in how a given part of the brain, and the whole body as such, determine the cognition of human beings and animals. Cognitive abilities are studied by neuroscientists investigating different groups of people. For example, neuroscientific research focuses on the cognitive abilities of people following brain injury or who are experiencing neurodegenerative diseases. Such interest in cognitive disorders leads to a better understanding of the needs and possibilities of patients suffering from, for example, Alzheimer disease and other forms of senile dementia. Another application of neuroscientific investigation for cognitive purposes is seen in the fields of management and marketing. The aim of neuromanagement is to check the cognitive side of leadership, studying the way a given style is perceived and how that style determines the perception of tasks. As far as neuromarketing is concerned, researchers are interested in which parts of the brain are responsible for purchasing behaviors and how a presented stimulus may determine the subject's reaction to the merchandise.

Technological Perspectives on Cognitive Science

Modern advancements in technology have led to the growing interest in *cognitive computing*, focusing on creating machines that possess human-like abilities. *Cognitive ergonomics* focuses on the relation between people and machines, studying, for example, the interaction between individuals and computers. Technological dimensions also consider differences connected with the growing popularity of the Internet in contemporary society. Cognitive science is also interested in studying the differences in the way people behave in online and offline environments. One field of investigation focuses on those who spend a large amount of time online and how this influences human relations and cognition as such. Another topic of research is the role of social networking tools in creating relationships and group participation.

Tools Used in Cognitive Science

Given its multidisciplinary nature, cognitive science relies on tools originating in different disciplines. Among the biggest donors is neuroscience since it offers both invasive and noninvasive apparatus to study how the brain and the nervous system function. One of the most complex machines and, at the same time, the one that offers many ways of looking at the tissue under study is *functional magnetic resonance imaging* (fMRI), which aims at measuring brain activity by observing changes in blood flow. The subject is placed within a special tube where anatomical scans are performed, followed by measuring the responses of subjects asked to perform certain activities. The closed and limited space within the apparatus is for some a source of discomfort; moreover, people with claustrophobia may be altogether unable to take part in this type of testing. Another inconvenience that is often reported by subjects is the level of noise generated by the apparatus (people may receive headphones that limit to some extent the level of noise). A key advantage of this method, however, is that it is noninvasive and does not carry risk for the subject. However, the use of the contrast agent that is injected to enhance the

visibility of the area under examination should be treated with great care if subjects report an allergy or are determined to be allergic to the agent. Another tool used in cognitive studies is *electroencephalography*, which records brain activity when the subject is sleeping, resting, or performing some activity, depending on the purpose of testing. Still another technique is *transcranial magnetic stimulation*, in which a wire coil is used to induce magnetic field. It offers the possibility of observing how the brain functions and serves as a tool in treating certain neurological problems. It should be mentioned that not only the brain itself is observed in researching the cognitive and behavioral processes of human beings. For example, *facial electromyography* is used to observe facial muscle performance and record the emotions of subjects to presented stimuli. Another technique, that of measuring *galvanic skin response* (skin conductance), concerns the electrical features of the skin and is used to study the subject's responses to presented stimuli. It should be mentioned, however, that not all tools are noninvasive, and sometimes scientists must rely on the invasive ones. For example, *positron emission tomography* makes use of radioactive isotopes to observe the blood flow in the brain during cognitive tasks. Another technique, *X-ray computed tomography* (X-ray CT), relies on radiation and a contrast medium and, consequently, cannot be retaken many times. Apart from such neuroscientific apparatus, cognitive science also benefits from the tools used in psychological and sociological testing. Thus, surveys, questionnaires, and interviews may be applied in cognitive studies. As far as psychology software tools are concerned, one such program is E-Prime, aimed at testing one's response to the offered stimulus, such as reaction time. The selection of tools varies with the type of experiment. For example, research focused on quantitative methods relies on statistical tools. Linguistic methodologies may also be applied successfully in cognitive science. For example, the techniques characteristic of *critical discourse analysis* can be used to check the way a given message is understood and perceived by the target audience. Such linguistic analyses facilitate the understanding of the role of mass media in creating texts appropriate to the recipients' cognitive abilities and expectations.

Ethics and Future of Cognitive Science

It is important to note that conducting research in cognitive science also raises ethical issues with respect to experimental protocols when conducting experiments and data management. First of all, individuals taking part in experiments should be informed about the purpose of the study and the further application of acquired data. As far as the data itself is concerned, attention should be paid to the way it is gathered and stored.

The future of cognitive science depends on progress in other fields of study. For example, new methods created by linguists to investigate the ways in which people communicate may also determine advancements in cognitive studies. As in the case of other fields of investigation, technological advancements also shape the progress of cognitive science. Technology in that case should be analyzed from at least two perspectives. One of them is treating technology as the factor of development, offering new tools of creating, gathering, and disseminating knowledge. In that case, the newest technological achievements influence the level of sophistication characterizing the methods of investigating cognitive notions. For example, the development of such neuroscientific tools as fMRI or galvanic skin response offers a new way of studying human cognition. It can be envisaged that together with the development of neuroscientific tools, cognitive science will also offer new ways of understanding the way living entities behave. It can also be predicted that the interest in cognitive science will be intense within both medical and commercial domains. Thus, cognitive studies, together with accompanying disciplines, will become even more central in efforts to decode human behavior and its implications.

Magdalena Bielenia-Grajewska

See also Actor–Network Theory; Linguistics, Contemporary; Neuroscience

Further Readings

Bermúdez, J. L. (2014). *Cognitive science: An introduction to the science of the mind*. Cambridge University Press.

- Bielenia-Grajewska, M. (2011). A potential application of actor network theory in organizational studies: The company as an ecosystem and its power relations from the ANT perspective. In A. Tatnall (Ed.), *Actor-network theory and technology innovation: Advancements and new concepts*. IGI, 247–258.
- Bielenia-Grajewska, M. (2013). International neuromanagement: Deconstructing international management education with neuroscience. In D. Tsang, H. H. Kazeroony, & G. Ellis (Eds.), *The Routledge companion to international management education* (pp. 358–373). Routledge.
- Bielenia-Grajewska, M. (2013). Schemas theory. In E. H. Kessler (Ed.), *Encyclopedia of management theory* (pp. 676–678). SAGE. <http://dx.doi.org/10.4135/9781452276090.n210>
- Bielenia-Grajewska, M. (2015). Company linguistic identity and its metaphorical dimensions: Purchasers, personnel and products through the perspective of metaphors. In N. Holden, S. Michailova, & S. Tietze (Eds.), *The Routledge companion to cross-cultural management* (pp. 170–179). Routledge.
- Bielenia-Grajewska, M. (2015). Neuroscience and learning. In R. Gunstone (Ed.), *Encyclopedia of science education* (pp. 698–704). SAGE.
- Bielenia-Grajewska, M., & Scarlett, G. (2015). Dynamic and relational systems theories. In G. W. Scarlett (Ed.), *The SAGE encyclopedia of classroom management* (pp. 397–443). SAGE.
- Eicher, J. (1996). Cognitive management: Managing your organization's mind. *Nonprofit Management Leadership*, 35, 12–15. <http://dx.doi.org/10.1002/pfi.4170350705>
- Friedenberg, J., & Silverman, G. (2012). *Cognitive science: An introduction to the study of mind*. SAGE.
- Reyna, V. F. (2012). Risk perception and communication in vaccination decisions: A fuzzy-trace theory approach. *Vaccine*, 30(25), 3790–3797.
- Stillings, N. A., Weisler, S. E., Chase, C. H., Feinstein, M. H., Garfield, J. L. & Rissland, E. L. (1995). *Cognitive science: An introduction*. MIT Press.

COMMUNICATION THEORY, TECHNICAL OVERVIEW

The terms *communication theory* (CT) and *information theory* (IT) are used interchangeably as synonyms in popular discourse. There are historical reasons for this that are easily explained. Also,

CT is often considered the modern derivation of *rhetoric*, dating back to Aristotle and Cicero. “Classical rhetoric,” however, is not the humanistic version of CT today, although *tropes of speech* are used as *logic models* for theory construction. CT currently refers to *communicology*, defined as the science of human communication using methodologies based upon the *normative logics* of human speech and gesture. CT is the *eidetic science of meaning* (semiotic signification, historically called *symbol theory* or *symbolology*). CT is a *binary analogue logic* that constitutes *certainly* (i.e., possibility differentiation) by combination judgments, simply formulated as a “choice of context.” CT is strictly formulated in logic as “the choice of a context by a relation which entails a choice by correlation” with the logic formalization: {Both [Both/And] And [Either/Or]}. Thus, CT explains a *message* that must express its own *code* simultaneously. Animals and machines cannot do this.

In simple terms, this logic formulation of communication as an *eidetic science* captures (*capta*) the basic facts of *human consciousness and comportment* (language) that operate at three code levels of consciousness unique to human beings. Human *codes* are (1) *synergistic* (qualitative rule: the whole is greater than the sum of the parts; Roman Jakobson's thesis that *metaphor* always dominates *metonymy*) and (2) *synesthetic* (quantitative rule: all human body senses are integrated simultaneously as one phenomenon; the thesis that *synecdoche* always dominates *simile/irony*) at three logical levels of expression and perception:

Affection or emotion. [pathos; *capta*]

Cognition or thought. [logos; *data*]

Conation or purposeful action. [*ethos*; *acta*]

The medieval philosophers use the respective shorthand terms *capta* (that which is *taken*; “awareness”), *data* (that which is *given*; “awareness of awareness”), and *acta* (that which is *done*; representation of awareness of awareness). Our contemporary terminology is *discovery*, *explanation*, and *explication* (or interpretation). The combined codes were *verba* (that which is *said*; *gērys* “communication tone”; see Figure 6).

By comparison, IT is the *empirical science of signal theory* (numerical algorithms) used in

electricity-based machines like computers, the name of which is *signal theory* or *informatics*. James Gleick offers a comprehensive social history of the information engineering concepts as they emerged in the United States in the 1950s up until the 1960s when, according to Gleick, in the social sciences, the direct influence of information theorists had passed its peak. Its specialized mathematics, he notes, has less and less to contribute to psychology and more and more to computer science.

Here, “psychology” (historically with Karl Bühler) is the example of the human sciences that advanced CT as prior to IT in theory construction. The critiques of Claude Shannon’s IT were made by Margaret Mead, Gregory Bateson, and Norbert Wiener (later by Jürgen Ruesch and Roman Jakobson) at the Macy Foundation Conference on Cybernetics held March 22–23, 1950, in New York City. Note that the “father of cybernetics,” Norbert Wiener, studied logic systems in 1914 at Göttingen University, Germany, with Edmund Husserl. IT is a *digital logic* that reduces uncertainty (i.e., constitutes *probability* differentiation) by exclusion judgments, simply formulated as a “Context of Choice.” IT is strictly formulated in logic as “the context of a choice by correlation which entails a context by relation” with the logic formalization: {Either [Either/Or] Or [Both/And]}. Machines are *spatially* limited to this logic, while animals are *temporally* limited by the logic. Thus, IT explains a *code* that must express a probability for all messages in the system (mechanical or biological).

In nonmathematical terms (“information” is an *algorithm*)—that is, in logical terms—the IT formulation gives (*data*) the basic quantitative procedure of *exclusion* whereby phenomena are negatively defined (*null hypothesis*) according to the presence [1] or absence [0] of an electrical signal in a digital system [0,1]. What is *not* the case [0] gives you a *probability guess* at what is the case [1]. If you can find one case to contradict the null hypothesis, then you have a correct judgment by probability. IT assumes a *closed system* (context) in which all choices are made. There are two important system facts here: (1) *Meaning* is not part of the theory and (2) the chosen message is *unknown*.

So, a machine or an animal organism is a closed system and probability of choice is *either* inside it or not. Which choice you make does not matter,

since either choice gives the same “information”; if, for example, you want a candy-bar from a machine, the machine signs reads “Insert \$1,” and you make a decision, or not, to insert a dollar. If you insert money, you may *either* get your candy or not [meaning is not in the system], the probability is the same; the only sure action is not to insert the dollar (not choosing the closed system). But you can take a chance by *both* inserting your dollar *and* hoping to get the candy bar [the message is uncertain]. IT always reduces the uncertainty of *how* you are choosing, but it never specifies *what* is chosen (either candy or no candy) among the possibilities! Machines are closed *spatial* systems; animals are closed *temporal* systems (information is limited to one generation). Digital systems (either/or = closed) can never operate *simultaneously* with analogue systems (both/and = open) because the digital operates at one logical level only. Only human beings can simultaneously operate an analogue of *three logical levels* by (1) simultaneously *both* remembering (2) what was chosen and not chosen [both/and] *and* (3) deciding which choice to act upon and which to ignore [either/or]. Remember that at stage (3), two choices are forever possible, even *after* a choice has been made. Humans call it “changing my mind,” that is, a “yes” and “no” becomes “maybe”!

Suffice it to say, much of the terminology confusion of CT and IT comes from the lack of careful research into stages of *theory construction* that emerged in 19th-century Europe, especially Germany where a distinction is drawn between the *human sciences* (*Geisteswissenschaften*), often called the “humanities” in the United States and the “eidetic sciences” in Europe, versus, the *natural* (physical) *sciences* (*Naturwissenschaften*). Keep in mind that in European history, all subjects studied systematically at the universities were “sciences” (especially literature) in contradistinction to theology and art. The same sort of name confusion is true for the 20th-century United States, where the distinction between “qualitative” and “quantitative” *methods* is equally confused by both sides of the debate over “two cultures” (human vs. nature), i.e., “sciences” versus the “humanities.”

To resolve these confusions, it is useful to review the standard framework for types of (1) *theory*, (2) *methods*, and (3) *models* (conditions of

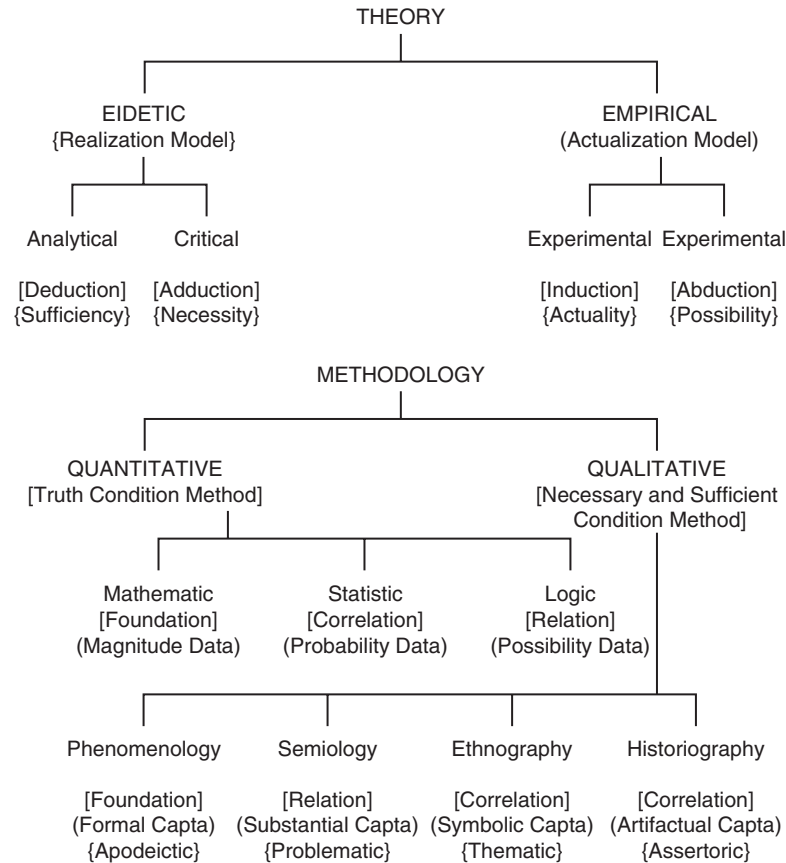


Figure 1 Theory and Methodology Chart

Source: Lanigan (1992, p. 207).

interpretation for validity and reliability). Then, a brief review of historically significant *communication models* will help orient us, although it is necessary to show the humanistic correction in Claude Shannon's informatics model. Last, a summary table formulates the present combined heritage of the reviewed theories and models as they are in use today under the disciplinary names "communicology," "speech communication," and "discourse communication."

Theory and Model Construction

Theory is a set of propositions; in other words, statements expressing the complete normative criteria for choosing contexts of judgment constituting a set of *process rules* (Figure 1). Where a theory is applied methodologically to a specific subject matter, it is called a *model*. There are two basic ways to construct a theory. *Eidetic theories* use

human *thinking* as a conceptual or ideational process to make sense of, or be conscious of, a phenomenon that *can exist*. Such a phenomenon is a *realization* (it can be made into subject and object) or an *idealization* (no object is possible yet). According to the American logician and semiotician Charles S. Peirce, *eidetic theory* is either *analytical* or *critical*.

Analytic Theory

Analytic theory uses a *deductive logic* [Rule + Case = Result]. In the *same* context, [Rule] two phenomena [Case] are *internally* compared [Result]. This analysis requires *sufficiency* as a standard of judgment. *Sufficient condition* in *logic* means that you can find at least one example of the phenomenon, but the example can be a *mentifact* (real) or an *artifact* (actual). *Analytical models* of theory construction are *validity tested*

reflexively by *critical* models (an internal test) and by *experiential* models (an external test).

Keep in mind that two *different* “phenomena” may be the *same* phenomenon at different times/ places (Roman Jakobson’s logic of *distinctive* and *redundancy* features in CT). This is a key point in the study of discursive phenomena, especially with regard to Alfred Schutz’s temporal model of *interpersonal communication*: (1) *Predecessors* are persons who never share time/space, for example, you and your great-grandparents, (2) *consociates* (associates) are persons who do share time/space, e.g., you and your classmates, (3) *contemporaries* are persons who share time, but not space, e.g., you and the president of France, and (4) *successors* are persons who do not share time/space but can, such as you and your great-grandchildren. Another applied example is Margaret Mead’s model of *intergenerational communication as culture*: (1) *Postfigurative cultures* are those where children learn from their forebears [= Data], (2) *cofigurative cultures* where peers learn from peers [= Capta], and (3) *prefigurative cultures* where adults learn from children [= Acta]. All cultures progress through these stages as more and more generations combine their learning into a *logic of practice* [= Verba].

Using the logic of the philosopher Immanuel Kant, *analysis* is the division of a *whole into parts* (synecdoche) or a *substance into attributes* (metonymy); validity is tested by using the reverse process of critical *synthesis*. IT uses this logic (either/or choices) as a *reliability condition*.

Critical Theory

Critical theory uses an *adductive logic* [Rule + Result = Case (universal; *a priori*)]. In *different* contexts [Rule], an *external* comparison [Result] establishes the identity of two phenomena [Case]. Criticism requires *necessity* as a standard of judgment; that is, the eidetic claim that the comparison holds true in all *types* of the case (universal; no exceptions, all *tokens* fit all *types*) and that the *truth* of the claim (*tonality*) is obvious *before* it is experienced (*a priori*). *Necessary condition* in logic means the criterion by which the *realization* of one phenomenon (e.g., playing a *game*) is constituted, validated, and confirmed by the realization of *another* phenomenon (knowing or not

knowing the *rules* of the game). *Critical models* of theory construction are *validity tested* reflexively by *analytical* models (an internal test) and by *experimental* models (an external test). Again following Immanuel Kant, *synthesis* (criticism) is the addition of parts into a whole or attributes into a substance. CT uses this logic (both/and choices) as a *reliability condition*. In advanced theory construction (Figure 2), CT logic entails IT logic producing *metaphor* (substance/whole) and *simile/irony* (part/attribute). Note that this entailment is a strictly human capacity because *human consciousness* remembers *both* all choices made *and* all choice *not* made simultaneously. Machines (spatial limit) and animals (temporal limit) cannot do this; therefore, IT *never* entails CT; that is, a computer (or even group of computers) can be either digital (IT) or binary analogue (CT), but not both.

Empirical Theory

Likewise, according to Charles S. Peirce, *empirical theory* is either *experimental* or *experiential*.

Experimental Theory

Experimental theory uses an *inductive logic* [Case + Result = Rule]. With two phenomena [Case], an internal comparison [Result] establishes the same context [Rule]. Experimentation requires *actuality* as a standard of judgment. Actuality is the empirical presence/absence of a phenomenon, which validates the presence/absence of another phenomenon. An experiment requires *actuality as a function of probability* as a standard of judgment; that is, the empirical claim that the comparison holds true for this and subsequent cases (particular *tokens* and *types* cannot be distinguished) and that the truth claim (*tonality*) is dependent on the case(s) chosen. *Experimental models* of theory construction are *validity tested* reflexively by *experiential* models (an internal test) and by *critical* models (an external test). *Experimentation* is a function of IT: probability can make you more and more certain of *how* you are choosing, but it never tells you (defines) *what* you are choosing or have chosen. The usual example here is a *dictionary*: you look up the *signification* of a *word* only to be referred to other words that you

LOGIC: Communicology vs. Informatics

HandChart © 2018, Richard L. Lanigan - International Communicology Institute

1 Communicology Aspects of Communication and Information Theories

Communicology Theorem: Communication Theory Entails {is prior to} **Information Theory Plato:** Collection [synagōgē] must precede a Division / *diáiresis* }. **Edmund Husserl:** *Noesis* precedes *Noema*.

Symbology > *Communication Theory Rule*: {Both [Both/And] And [Either/Or]}

Informatics > *Information Theory Rule*: {Either [Both/And] Or [Either/Or]}

Communication Logic Rule:	Abduction (Particular; <i>a posteriori</i>)	{Rule+ Result = Case}
	Adduction (Universal; <i>a priori</i>)	{Rule + Result = Case}
Information Logic Rule:	Deduction	{Rule + Case = Result}
	Induction	{Case + Result = Rule}

2 Communication {Symbol} Theory → → Information {Signal} Theory

1. Def: Constitution of Intentionality Intentionality as present: "subjective"

= Perception of Expression (*Noesis*; Sign)

CYBERNETICS:

[*kybernētēs*: steersman, governor;
To Guide Human Comportment]

THIRD ORDER ("Human Choice")

[*Organism* Structure > Behavior> Development]

(*Semiotic Phenomenology*)

(*Phenomenological Semiotics*)

2. Consciousness-of-Experience

(*Medium and Meaning [Code]:*
understanding judgment/ comportment)

{ **Logic Step One: Rule ↓** }

3. [counterpoint: construction of certainty]

4. Choice of Context (Metaphor/ Simile)

{ Eidetic: Jakobson's "*Distinctive Feature*" }

5. Analogue Logic (**Both/ And**)

{ *Apposition* > *Triadic Analogue* > *Add* }

6. Necessary Condition as Hypothesis [Motivation]

7. Judgment by Theory Definition

[Genus → *Differentia*] (tropic logic = rhetoric)

{ **Logic Step Two: + Result ↓** }

8. Linguistic Theory [Hjelmslev's "Structure"]

9. Key Concept: CODE [codex: to choose]

{ **Logic Step Three: = Case** }

10. Theory Paradigm:

Contingency of Choice {Addition}

Primary Dynamic Logic:

John von Neuman (1944) Norbert Wiener (1948, 1950)

Secondary Cultural Application:

Jurgen Ruesch & Gregory Bateson (1951)

Ernst Cassirer (1923; 1953)

Roman Jakobson (1958 , 1960)

Hubert G. Alexander (1967)

Hubert L. Dreyfus (1972; 1992)

Anthony Wilden (1972; 1987)

Richard L.Lanigan (1988; 1992)

Irl Carter (2011)

1. [counterpoint: re-constitution of intentionality as absent: "objective"]

= Perception of Object (*Noema*; *Symbol*)

FIRST ORDER ("Machine Regulation")

[*System* Co-ordination > Regulation > Control]

(*Command and Control Systems Theory*)

SECOND ORDER ("Social Organization")

[*System Integration* > Learning > Evolution]

(*Socio-Cybernetics*; *Bio-Semiotics*) .

(*Semio-Cy be me tics*)

2. Experience-of-Consciousness

(*Channel and Signification [Message]:*
explaining decision/behavior)

{ **Rule ↓** }

3. Def: Reduction of Uncertainty

4. Context of Choice (Metonymy/Synecdoche)

{ Empirical: Jakobson's "*Redundancy Feature*" }

5. Digital Logic (**Either/ Or**)

{*Opposition* > *Binary Digital* > *Delete*}

6. Sufficient Condition as Hypostatization [Causality]

7. Judgment by Method Rule

[Genus → *Definiendum*]

(figurative logic = grammar) {+Case ↓ }

8. Mathematical Theory [Hjelmslev's "Form"]

9. Key Concept: MESSAGE [mittere: to send] {=Result}

10. Theory Paradigm:

Confinement of Context {Subtraction}

Primary Static Logic:

Claude E. Shannon (1948)

Warren Weaver (1949)

Secondary Social Application:

W. Ross Ashby (1956)

B. F. Skinner (1957)

P. Watzlawick, Beavin, Jackson (1967)

Niklas Luhmann (1970)

Ralph E. Anderson & Irl Carter (1974)

Heinz von Foerster (1949; 2002);

Klaus Krippendorff (1979; 2009)

Figure 2 Comparison of Communication Theory (Human) and Information Theory (Signal)

do not know. The dictionary is a *closed system*, which you already know, especially if you do not know how to *spell* the word (spelling is another closed system) for which you are searching.

Experiential Theory

Experiential theory uses *abductive logic* [Rule + Result = Case (particular; *a posteriori*). In the *same* context [Rule], an internal comparison [Result] establishes the identity of two phenomena [Case]. Experience requires *possibility* as a standard of judgment; that is, the empirical claim that the comparison holds true for this *type* of case (particular; all *tokens* fit the definition of *type*) and the *truth* of the claim (*tonality*) is obvious after it is experienced (*a posteriori*). *Experiential models* of theory construction are *validity tested* by *experimental* models (an internal test) and by *analytical* models (an external test). Experience is a function of CT; in other words, possibility can make you more and more certain of *what* you are choosing, but it never tells you (defines) *how* you are choosing or have chosen. The parallel example is an *encyclopedia*, where you look up the *meaning* of an *idea* that is fully explained by the entry. The encyclopedia is an *open system*, which you do not expect; there are more and more related *contextual* entries you can read (what you are doing right now!).

Method and Methodology

A *method* is a set of practices that are the actions a researcher uses to express the complete normative context (usually eidetic) for making choices among phenomena. *Methodology* is the creation of a *system* of methods for the constitution or production of *results* in logic. Where the context is limited (incomplete), it constitutes a *procedure*. Where there is only one known context, it is an *exemplar*. If you build a scaled physical version of the exemplar, it is called a *prototype*; for example, *writing* is a prototype of speech. All methodologies are *both quantitative and qualitative* as measure of space and time. For example, think of the verb tenses in the language you speak, the *grammar system* specifies space *quantities* (first, second, third, etc. person) and time *qualities* (present, past, future, gerund, etc.).

Quantitative methodology is the *practice* of using a *truth condition method* in order to compare and contrast choices made in a given context of magnitude, probability, or possibility. This is to say, all choices are true, false, or indeterminate (contingent). There are *three basic quantitative methods*: (1) *Mathematics* measures *magnitudes*; that is, differences by *degree*, such as a small versus a large, usually from a known or assumed reference point (a *foundation*) to a new point of reference (that which is “grounded”). Popular versions of this methodology are currently called “grounded theory,” although it is not a theory, but a mere *procedure*. In the human science context, the speech tropes of *metonymy* (substance and attribute comparison) and *synecdoche* (part/whole comparison) exemplify a *magnitude* measure in discourse. (2) *Statistics* measures *probability* (difference of kind and degree, e.g., two small vs. one large), usually from two assumed reference points (“dependent variable”) to two new points (“independent variable”) as a *correlation* (an either/or *disjunction*). A language example is the adjective system: “good, better, best”; such measures are often *not* logical, for example, “perfect, more perfect, most perfect.” The tropes of *simile* (positive part/attribute comparison) and *irony* (negative part/attribute comparison) exemplify a *probability* valence measure in discourse. (3) *Logics* measure *possibility* (difference by kind, e.g., an edible orange vs. a rotten apple), usually from a known or assumed reference point (a *class* or *type* such as “edible” or “rotten”) to a new point (that which is a *member* or *token* of a class such as “orange” or “apple”). The trope of *metaphor* (substance/whole comparison) is a discourse example of such a *relation* (a both/and *conjunction*).

Qualitative methodology is the practice of using a *necessary and/or sufficient condition* in order to compare and contrast choices made within a taken context of artifactual, symbolic, substantial, or formal *capta*. A *necessary condition* is a logical criterion by which the *realization* of one phenomenon is constituted, validated, and confirmed by the *realization* of another phenomenon; for example, in CT, any choice provides its own *context of meaning*; the choice is a necessary condition for the meaning. A *sufficient condition* is the logical criterion by which the *actualization* of one phenomenon is constituted, validated, and

confirmed by the *actualization* of another phenomenon; for example, in IT, either a “yes” or “no” choice provides the *same information*, or in other words, it affirms the *existence of the context*; each choice is a sufficient condition of the other choice.

There are four basic qualitative methodologies:

- (1) *Phenomenology* measures a narrative discourse as *true* by necessary condition (*apodeictic*) using formal capta (understanding) with a qualitative logic, rhetoric, or tropic of *foundation* (if/then logic). By tradition (*Ideographic Science* = “individualizing” judgment) in the discipline of *philosophy*, an *apodeictic* discourse is *critical*; therefore, validity is obtained in the application of an *analytical* review.
- (2) *Semiology* measures a narrative as *intuitive* by necessary condition (*problematic*) using substantial capta with a qualitative logic, rhetoric, or tropic of *relation* (both/and logic). Also by tradition (*nomothetic science* = “generalizing” judgment) in the discipline of *linguistics*, a *problematic* discourse is *experimental*; therefore, validity is obtained in the application of an *experiential* review. *Semiotics* is a synonym for “communication,” where communication includes the study of human, animal, and machine “languages” as icon, index, or symbol code systems.
- (3) *Ethnography* measures a *thematic* discourse using symbolic, mentifactual capta (knowledge) with a qualitative logic, rhetoric, or tropic of *correlation* (either/or logic). Traditionally (*Kulturwissenschaft* = cultural practice) in the discipline of *anthropology*, a thematic narrative is *experiential*; therefore, validity is obtained in the application of an *experimental* review. Ethnography is the systematic study of human cultures with particular attention to “space binding” relationships of a society (*Gesellschaft*) and culture (*Weltanschauung*).
- (4) *Historiography* measures an *assertoric* narrative (observation) using artifactual capta with a qualitative logic, rhetoric, or tropic of *correlation* (either/or logic). By tradition (*Geistgeschichte* = historiography) in the discipline of *history*, an *assertoric* discourse is *analytical*; therefore, validity is obtained in the application of a *critical* review. Historiography

is the systemic study of human cultures with particular attention to “time binding” relationships of an individual (*Lebenswelt*) and community (*Gemeinschaft*), for example, language, kinship, and ritual.

Communication Models: Karl Bühler

The first formal model of human communication was suggested by Karl Bühler in his 1934 book *Sprachtheorie* (Theory of Speech), where he depicts the semiotic process of human communication as an interpersonal interaction including the complex linguistic and psychological context. Bühler is correcting the semiotic phenomenology of Edmund Husserl and laying the procedural groundwork for Roman Jakobson’s later research using Russian poetry as a prototype. Unfortunately, the book was not translated into English until 1982; hence, it did not have much of an early influence in the United States, despite the fact that Bühler emigrated to the United States from Germany in 1940.

The model shown in Figure 3 depicts the acoustic phenomena of speech communication as a complex *sign* [s] symbolizing the *apperception* process by which the language sign displays the *apposition* of (1) a *symbol* of the objects and states of affairs or *representation* of the referent, (2) a *symptom* dependent as an index of the

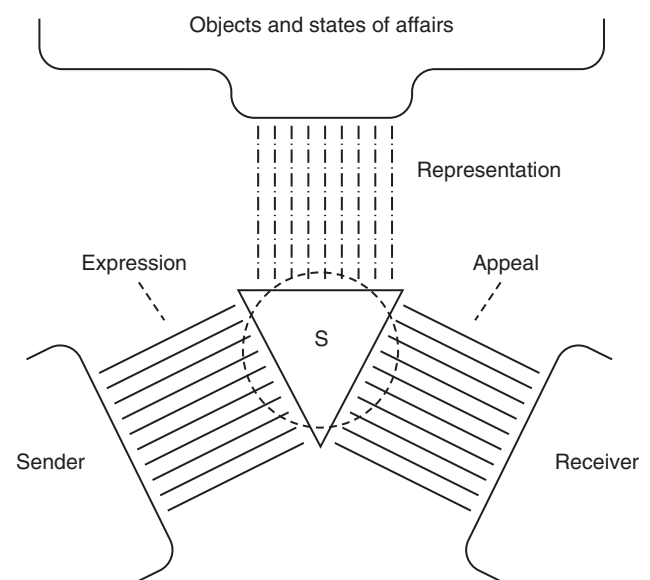


Figure 3 Bühler Communication Model

sender or *subject of expression*, and (3) a verbal *signal* for the interpretation of the receiver or an *appeal* to the *addressee's* perception. All the parallel lines in the diagram illustrate the complex *semantic* functions at work in the creation of communicative *meaning*.

Communication Models: Claude Shannon

In 1948, Claude Shannon published a technical paper as part of his job at the Bell Telephone Laboratories entitled "A Mathematical Theory of Communication." The paper was of interest primarily to electrical engineers working on telephones and two-way radios, one of whom was Warren Weaver. Weaver wrote a popular explanation of the theory for the magazine *Scientific American* in 1949. In the article and subsequent reprints, he argued quite *incorrectly* that the

general model depicted in Figure 4 could be generalized from *closed* machine systems to human communication as an *open* system. Unfortunately, Weaver had no knowledge of logic or linguistics and was not at the Macy Conference in 1950 where his errors were corrected. His unjustified generalization started a confusion that is still with us today in which IT (Signal Theory and Informatics) is confused with CT (Symbol Theory, Linguistics, and Rhetoric). Fortunately, Shannon corrected Weaver by noting that for *human communication symbols* using an electrical machine, the process shown in Figure 5 would be necessary. Here, a human observer using human *language* is always part of the construction and corrective maintenance of the communication *process*. There are two important IT facts here: (1) *Meaning* is not part of the theory and (2) the chosen message is *unknown*. Keep in mind as well that Shannon's

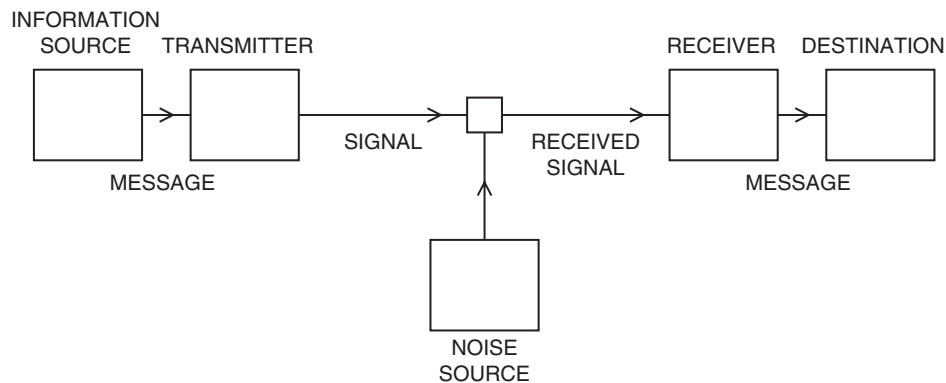


Figure 4 Shannon Signal Model of Informatics

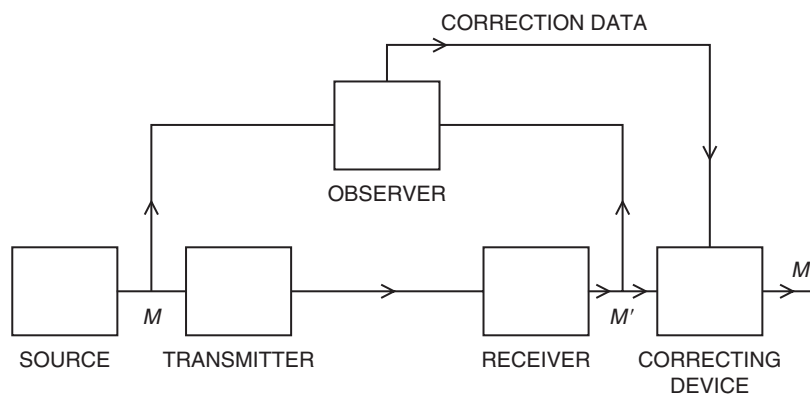


Figure 5 Shannon Observer Model of Informatics

model is *linear* (one-dimension, directional action with control = cybernetics) and does not account for the complexity of *curvilinear* human interaction (three dimension, multidirectional transaction = semiotics).

Communication Models: Ruesch and Bateson

As an immediate correction of the IT and CT confusion, Jürgen Ruesch and Gregory Bateson published their historic book *Communication* in 1951. The volume immediately set down the conditions for a “human observer” model of communication (Figure 6). Note in fact that this book sets the basic divisions for research and teaching that

have become the main areas of the *human communication discipline*: (1) intrapersonal communication, (2) interpersonal communication, (3) group communication, and (4) cultural communication. As Ruesch and Bateson (2008) specify, “The focus of the human observer is not fixed; rather it has to be viewed as a fluctuating or oscillating phenomenon in which quick glances are taken rapidly at various levels and at various functions. Communication is an extremely dynamic phenomenon with a rapid rate of change of levels and functions, which range from evaluation to transmission and conduction” (p. 274).

The famous “Table D” from this book defined the research agenda of the discipline (note that

JÜRGEN RUESCH AND GREGORY BATESON DISCOURSE OBSERVER MODEL OF COMMUNICOLOGY (1951)

Handchart © 2019, Richard L. Lanigan, Ph. D.

THE LEVELS OF COMMUNICATION

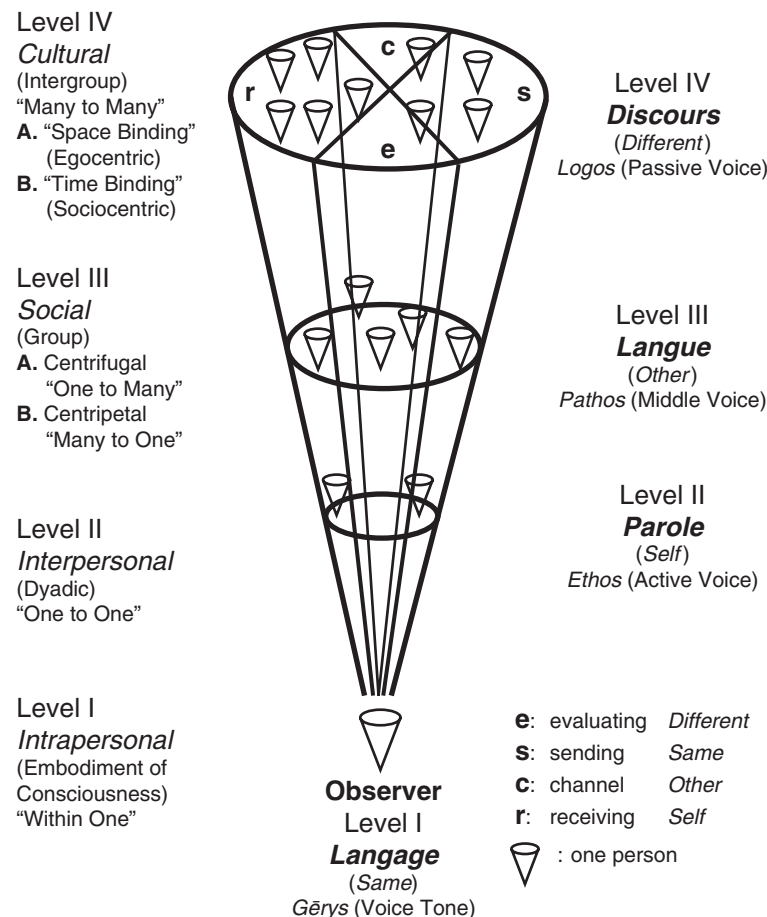


Figure 6 Ruesch and Bateson Observer Model of Communication

Table 1 summarizes an updated version). We need to recognize that Jürgen Ruesch is the virtual founder of the modern human science discipline of communicology with his collected work appearing in *Semiotic Approaches to Human Relations* (1972) and his *Knowledge in Action* (1975), which is the foundational research devoted to “communication, social operations, and management.”

Communication Models: Roman Jakobson

By 1958, CT was fully integrated with linguistics and semiotics, and Karl Bühler’s work was emerging indirectly in that of Roman Jakobson. A linguist who wrote in eight languages, Jakobson conceives of communication as the unifying theory that connects all the levels of the human sciences, both in theory and methodology. His

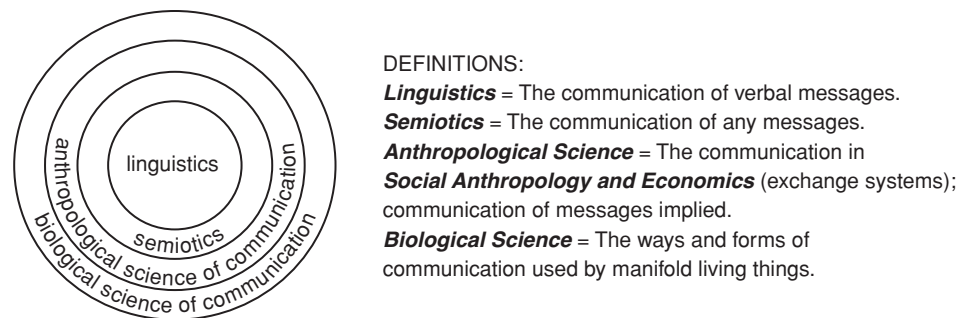


Figure 7 Roman Jakobson Human Sciences Model

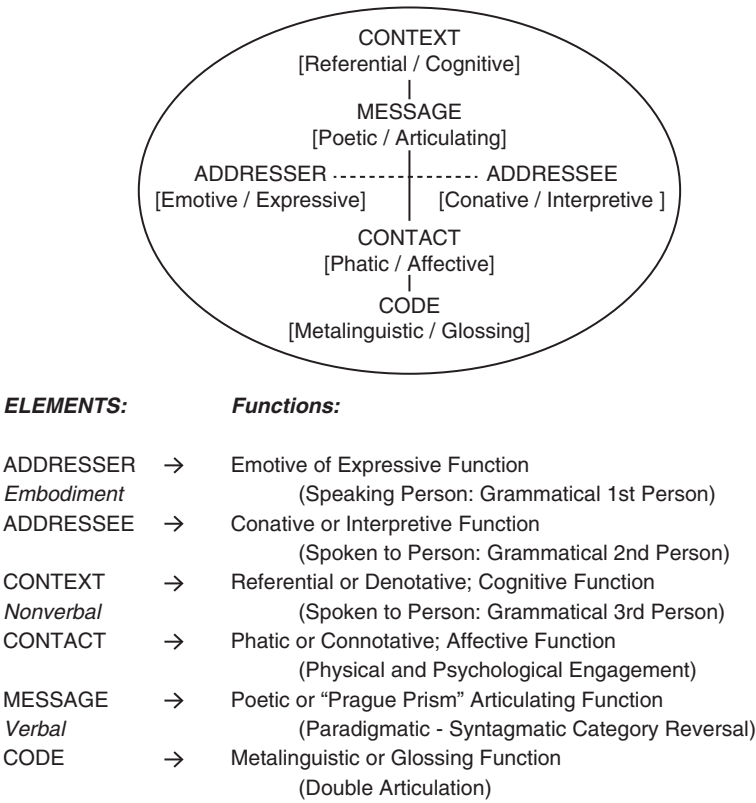


Figure 8 Roman Jakobson Model of Communication

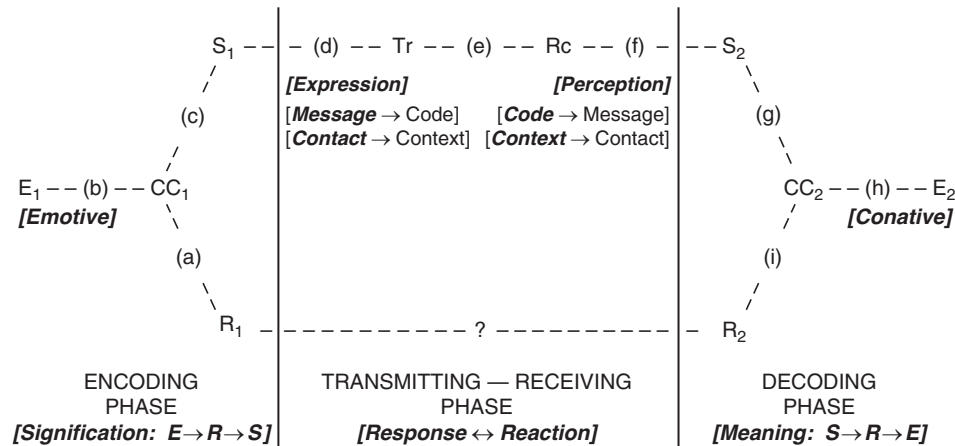


Figure 9 Hubert Alexander Model of Communication

Key to symbolic notation: E_1 = The Background Experience and Attitude of the Communicator; CC_1 = The Concept of the Communicator; S_1 = The Symbol(s) Used by the Communicator; R_1 = The Referent(s) as Perceived or Imagined by the Communicator; S_2 = The Symbol(s) as Understood by the Communicatee; R_2 = The Referent(s) as Perceived or Imagined by the Communicatee; CC_2 = The Concept of the Communicatee; E_2 = The Background Experience and Attitude of the Communicatee; Tr = The Transmitting Device, Mechanism, Medium of Expression; Rc = The Receiving Device, Mechanism, Medium of Perception; ? = The Mediation Possibility of a Relationship as Response (Verbal), Reaction (Nonverbal); a—i = Specific Boundary Relationships That Are Necessary (validity) and Sufficient (reliability) Conditions in the Communicological Process

human science model is illustrated in Figure 7 along with the basic definitions of the hierarchy of development among the various eidetic and empirical science disciplines. He begins with *linguistics* as foundational moving logically outward to develop the general communication systems of all organic life forms.

Jakobson goes on to specify his model of human communication, as shown in Figure 8. The model consists of six *elements* of the communication process with six corresponding *functions* of transaction. While the model is illustrated in the procedural terms of dyadic interpersonal communication, it is critical to understand that the dyadic process can apply at *all four levels* of the Ruesch and Bateson model. This is to say, the model explicates all communicative levels from intrapersonal (e.g., one's own thinking) to interpersonal (e.g., conversation), to group (e.g., a work team), and to cultural (e.g., a global conference).

Communication Models: Hubert Griggs Alexander

By 1967, CT became a focal point in the discipline of *philosophy*, which today flourishes as the “philosophy of communication” in most of the

professional associations worldwide. As a student of Ernst Cassirer, Edward Sapir, and Wilbur Marshall Urban at Yale University, Hubert Griggs Alexander's work is the original classic contribution to explicating the *philosophic* concern with how human thinking operates as *rationality* through the interactive agency of language and logic. Figure 9 reproduces Alexander's model along with some insertions [in brackets] to show the correlation with Roman Jakobson's model.

While this model is a *prototype formalization* of the communication process, its simplicity in specifying the elements, functions, and stages of the process is a masterpiece of *procedural description*. Using the model allows a comprehensive description of each aspect of the *encoding* and *decoding* processes that govern communicative and semiotic *competence*. The model is most useful from the point of view that communication *failure points* can be located and described with great precision, that is, points at which *performance* misunderstanding displaces understanding. For ease of use, these process functions are presented in outline form. Outline lettering (a, b, c, etc.) corresponds to model lettering in Figure 9. Each example is a *prototype* of actual communication practice.

Table 1 Communication Theory Table of Elements, Functions, and Levels

<i>Elements →</i>	<i>Addresser</i>	<i>Context</i>	<i>Message</i>	<i>Contact</i>	<i>code</i>	<i>ADDRESSEE</i>
<i>Verbal →</i> <i>functions</i>	<i>Emotive</i>	<i>Referential</i>	<i>Poetic</i>	<i>Phatic</i>	<i>Metalinguistic</i>	<i>Conative</i>
<i>Nonverbal →</i> <i>functions</i>	<i>Vocalics</i>	<i>Proxemics and</i> <i>ocularics</i>	<i>Chronemics</i>	<i>Haptics</i>	<i>Kinesics</i>	<i>Olfactorics</i>
<i>Media network levels: Each level is a medium of communicology</i> (Jürgen Ruesch, Margaret Mead, and Gregory Bateson)						
① Intrapersonal	Embodiment: Self awareness, <i>synesthesia</i>	Content ordering: <i>Consciousness</i>	Store signification: “Within One”	Pre- consciousness: <i>Pre-reflectivity</i>	Synesthetic meaning	Memory: Synchronic (Here/Now) <i>Present</i> as past/future
② Interpersonal	Dyad: <i>Self as</i> other; relationship	Task ordering: <i>Experience</i>	Transmit meaning: “One to one”	Consciousness; <i>Reflectivity;</i> conscience	Cognitive meaning	History: Diachronic (There/Then) Past/future as present
③ Group	Community: <i>Other as self;</i> affinity; kinship	Group ordering: <i>Socialization</i>	Retrieve signification	Present: <i>Reflexivity;</i> altruism	Affective meaning	Consociates: Share space–time
① (Individual) Egocentric	Task group: <i>Identity by rule</i>	Aggregate parts	Centrifugal: “One to many”	Competition: Creates <i>agony</i>	Names create static categories	Primary roles: Leadership
② (Group) Sociocentric	Affiliation group: <i>Identity by role</i>	Organic whole	Centripetal: “Many to one”	Cooperation: Creates <i>harmony</i>	Names create dynamic relations	Secondary roles: Membership
④ Culture	Co-figurative: “Peers learn from peers”	Intergroup ordering: <i>Culturalization</i>	Evaluate, interpret meaning	Present: <i>Reversibility of choice</i>	Conative meaning	Contemporaries: Share time only
① Space (Position/ Place)	Post-figurative: “Children learn from forebears”	Place community: “Family” (Gemeinschaft)	Space binding Decode: O of E “Many to many” (“Home”)	Past: <i>Practice orientation</i>	Digital logic: <i>In-group vs. out-group</i>	Predecessors: Previous generations; ancestors
② Time (Moment/ Event)	Pre-figurative: “Adults learn from children”	Nonplace community (“Profession”) (Gesellschaft)	Time Binding Incode: O of A “Many to many” (“House”)	Future: <i>Practice orientation</i>	Analogue logic: Diffusion of innovations	Successors: Succeeding generation; progeny

Channel elements and functions: Each category is a *channel* of communicology

(Karl Bühler and Roman Jakobson)

© 2016 Richard L. Lanigan

❶ Failures in Conceiving and Encoding

(Signification = Language System; Message-to-Code).

- (a) **Referent:** Epistemic-Pragmatic Failure from forming the wrong idea by the Communicator may be due to:
 - [1] Unclear Perceptions (e.g., mistaking a “cat” for a “dog”).
 - [2] Improper Conceptualizing (e.g., mistaking a cardinal [“1”] for an ordinal [“first”] number).
 - [3] Wrong Application of a Properly Conceived Idea (e.g., taking Earth to be a perfect sphere).

- (b) **Experience:** Epistemic-Emotive Failure because of Preconceptions on the part of the Communicator may be due to:
 - [1] Lack of Appropriate Background Experience.
 - [2] Inadequate Education.

- (c) **Symbol:** Semiotic-Semantic Failure in Encoding (Using the Wrong Symbols) may be due to:
 - [1] Inadequate Knowledge of the Symbol System (e.g., saying “dog” when one means “cat”).
 - [2] Using Ambiguous Symbols, or Symbols that are apt to be unfamiliar to the Communicatee, either intentionally (deception) or not (mistake).

❷ Failures Due to Expression (Transmitting) and Perception (Receiving).

- (d) Physiological Failures of the Communicator (e.g., mis-articulation, stuttering).
- (e) Mechanical Failures in Transmitting and Receiving (e.g., mispronunciation, noise).
- (f) Physiological Failures of the Communicatee (e.g., being hard of hearing, unfamiliar accent).

❸ Failures in Decoding and Discovering the Correct Referent (Meaning = Speech System; Code-to-Message).

- (g) **Symbol:** Semiotic-Semantic Failure in Decoding (Using the Wrong Concept from the Symbols Received) may be due to:
 - [1] Inadequate Knowledge of the Symbol System Used by the Communicator (e.g., hearing a foreign or unfamiliar language).
 - [2] Not Recognizing All of the Implications of the Symbols Used (e.g., not detecting irony in the speaker’s tone of voice).

- (h) **Experience:** Epistemic-Emotive Failure because of Preconceptions on the Part of the Communicatee (e.g., having a misconception of geographical orientation, which leads to a failure to understand directions; or, having no experience with an object which is being described to one).
- (i) **Referent:** Epistemic-Pragmatic Failure in Finding the Correct Referent (e.g., failure to get from the correct, but too general, concept to a specific referent)

Communicology: Human CT

CT as a human science represents one example of complexity in theory construction with attendant methodological concerns. Figure 9 allows a comprehensive synthesis of many of the content and structural relationship issues that exist in the *discipline of communicology*. However, it is essential to remind ourselves that theory and methodology are options that may be combined, and are, to solve particular research problems. The key to understanding the “humanistic” influence on communication is that the correct name is *human science* and that *science* uses rules of validity and reliability that must be specified when research result claims are made. Historically, *human science* relies on *logic* to make eidetic and empirical claims. By comparison, the *physical sciences* rely on *mathematics or statistics* to make similar claims. The concepts (terms) of *eidetic* (human consciousness) and *empirical* (human experience) are badly misused by most people who have not been trained in theory construction in their particular discipline, largely because the *practice* of most research is driven by habitual *procedures* of unquestioned methodology. Table 1 *categorizes* the issues of *theory construction* that explicate and explain human beings communicating with human beings. Methodology is not illustrated because it is a secondary concern, since all methodologies are used by all disciplines in both the human and physical sciences.

Richard Leo Lanigan

See also Cybernetics, 20th Century; Linguistics, Historical; Logic, Formal and Informal; Logical Theory; Semantics (Scientific and Empirical); Theories, Semantic Conception of

Further Readings

- Alexander, H. G. (1998). Communication. In *The language and logic of philosophy* (pp. 11–32). University Press of America. (Original work published 1967)
- DeVito, J. A. (1978). *Communicology: An introduction to the study of communication*. Harper & Row.
- Fisher, H. (2009). *Self, logic, and figurative thinking*. Columbia University Press.
- Gleick, J. (2011). *The information: A history, a theory, a flood*. Pantheon.
- Holenstein, E. (1976). *Roman Jakobson's approach to language: Phenomenological structuralism* (C. Schelbert & T. Schelbert, Trans.). Indiana University Press.
- Lanigan, R. L. (Ed.) (1992). Communicology: An encyclopedic dictionary of the human science. In *The human science of communicology: A phenomenology of discourse in Foucault and Merleau-Ponty* (pp. 197–236). Duquesne University Press.
- Lanigan, R. L. (1997). Communicology. In L. Embree (Ed.), *The encyclopedia of phenomenology* (pp. 104–110). Kluwer Academic Publishers.
- Lanigan, R. L. (2010). The verbal and nonverbal codes of communicology: The foundation of interpersonal agency and efficacy. In I. E. Catt & D. Eicher-Catt (Eds.), *Communicology: The new science of embodied discourse* (pp. 102–128). Fairleigh Dickinson University Press.
- Lanigan, R. L. (2011). Defining the human sciences. In *Schutzian Research—A Yearbook of Lifeworldly Phenomenology and Qualitative Social Science*, 3, 9–11.
- Lanigan, R. L. (2013). Information theories. In P. Copley & P. Schulz (Eds.), *Theories and models of communication, vol. 1* (Handbooks of Communication Science, 22 vols., 2012–2019, pp. 59–83). De Gruyter Mouton.
- Lanigan, R. L. (2017). Ernst Cassirer's theory and application of communicology: From Husserl via Bühler to Jakobson (Special Issue: Cassirer on Communicology, ed. R. L. Lanigan) *American Journal of Semiotics*, 33(3/4), 181–231.
- Lanigan, R. L. (2018). Franklin H. Knowler (1901–1993) — A model for communicology: Edited reprint with an introduction by Richard L. Lanigan. *Review of Communication*, 18(4), 391–413.
- Ruesch, J. (1972). Synopsis of the theory of human communication. In *Semiotic approaches to human relations* (pp. 47–94). Mouton. (Original work published 1953)
- Ruesch, J., & Bateson, G. (2008).

Individual, group, and culture: A review of the theory of human communication [includes Table D]. In *Communication: The social matrix of psychiatry*. Transaction Publishers. (Original work published 1951)

COMPLEMENTARY MEDICINE

Complementary medicine is the provision of traditional remedies in concert with conventional, Western-based, scientifically informed medicine. Each type of practitioner may provide treatment in the form of substances to be consumed or applied to the body in some way or therapies to be enacted on the body or mind. However, in the case of traditional medicine, typically, the herbal remedies or behavioral modifications suggested by healers have not been tested rigorously by controlled clinical trials, investigating outcomes and risks. Instead, the practices are supported by beliefs, personal accounts, and hopefulness. This entry describes the origins and nature of complementary medicine, with particular emphasis on current practice in the United States.

The strong role of customary beliefs in traditional medicine and the variety of cultures around the world have resulted in many different traditional practices being included in the category of complementary medicine, with no overarching philosophy connecting these disparate treatment regimes. For example, among the Dogon, in southern Mali, childhood illness may well be regarded as the effect of witchcraft, supported by local cosmological understandings of the world, and treated with culturally specific substances and behaviors. These Dogon healing practices are different from those found among Maya Q'eqchi', in rural Guatemala, where there may be similar local reactions to childhood illness, but interpreted according to Q'eqchi' beliefs, and with different treatment protocols. In traditional medicine, there is no overarching system incorporating these diverse healing methods, as is the case with mainstream medical understanding, based on the scientific testing of theories, with ongoing evidence-based treatment improvements.

In addition to lacking scientific verification, traditional healing is characteristically more

broadly focused than the conventional medical concentration, which typically deals only with the body or the mind. Healers may also incorporate a spiritual element, often accompanied by attention to the patient's relationships with others, such as the family or the community or even the environment. This means that curing events may be viewed as religious ceremonies and may also involve the members of the extended family, sometimes accruing considerable financial cost and organizational challenge.

Traditional healers usually are not medical school graduates but instead have learned their techniques from other traditionalists. Curers provide their treatments in villages and towns around the world, especially in less-developed nations, where traditional belief systems are supported, such as Nigeria, Cambodia, or Guatemala, and increasingly in large cities with immigrant populations, in developed nations such as the United States and Canada.

In North America, many minority community members use their cultural healing practices in combination with mainstream, Western-based medicine. A smaller group of people rely solely on traditional medicine, and in these cases, individuals are employing alternative medicine rather than complementary medicine. People may shift between these strategies at different life stages, depending on their conceptualization of health, their understanding of the practices, and the type of ailment experienced.

In addition to shifts in individuals' use of health therapies throughout their lives, there are also advances in medical knowledge, which mean that the boundary between acceptable and unacceptable healing practices is in flux. Indeed, many formerly traditional practices are now accepted as standard medical techniques, such as the use of massage, which is now often combined with exercise at physiotherapy clinics treating back injuries.

Sometimes, people access complementary medicine by requesting referrals from their mainstream medical provider. This might be the case for back injuries, when a family doctor may suggest visiting a chiropractor. This type of referral may be covered by some types of health insurance.

At other times, people find complementary medicine through their own networks and may

begin activities such as meditation or consuming products such as megavitamins or herbal mixtures, based on recommendations from friends, local community members, internet reviews, or other sources. Some of these individuals may consult their medical doctor when deciding to try these treatments, while others just use their own judgment. In most of these cases, the complementary medicine accessed would not be covered by health insurance. Regardless of whether people have health insurance coverage, a wide variety of complementary medicine types are available in the United States.

Types of Complementary Medicine in the United States

The first Americans, namely, American Indians and Alaska Natives, lived very different lives, spoke very different languages, and had very different traditional medicines, one from another. Different tribal groups inhabited dissimilar regions, each with distinctive flora and fauna, which could be harvested and used for curing. Occasional accounts of aboriginal treatments were recorded in early histories. For example, in 1535, Jacques Cartier, while exploring what is present-day Quebec province, Canada, encouraged his sailors suffering from scurvy to consume an Iroquoian drink based on boiled coniferous needles, which contain ascorbic acid and which cured the underlying vitamin deficiency.

Despite a few early positive interactions between explorers and American Indians, with ongoing European contact and settlement, many aboriginal groups were separated from their land and lost their knowledge of local plants, the basis of customary healing lore. Although much has vanished, some groups have managed to retain traditional knowledge, and it is typically available to indigenous people through healers, some of whom may be found at health clinics or through community providers. Another source of traditional knowledge is through American Indian doctors working in clinics, who have reconnected with cultural beliefs to incorporate these in a complementary fashion in their contemporary medical practices. This was the case for Dr. Lori Arviso Alvord, who received her medical degree from Stanford University, becoming the first

surgical board–certified Navajo woman in 1994. Dr. Alvord works holistically with her Navajo patients, providing a Western medicine–based practice, incorporating traditional beliefs about balance and relationships, described in her book, *The Scalpel and the Silver Bear*.

Balance in individuals, families, and the world is important in Navajo cultural cosmology. Transition periods, or other times of imbalance, are treated through curing ceremonies involving chants, sand paintings, food, and other ritual activities to restore stability. People who have grown up in the remote reaches of Navajo Nation feel validated to visit a doctor who understands their foundational beliefs. However, this type of complementary practice remains atypical for American Indians and Alaska Natives because there is a paucity of indigenous doctors graduating from medical schools.

In recent history, Native American populations have been outnumbered by newer residents. Each of these different immigrant groups has imported their own beliefs and ideas about health, illness, and treatment to the United States. For example, in the 19th century, Chinese herbal shops were located in many urban centers in California and were heavily utilized by local Chinatown residents. At the time, outsiders did not typically frequent these shops, given the high levels of discrimination against Chinese immigrants. More recently, with improved ties between China and the United States and decreased stereotyping of the California communities, some traditional practices, such as acupuncture, have increased in popularity, becoming well known among the users of complementary medicine.

Even more widely accepted than acupuncture is yoga, which is originated in ancient India. This practice involves moving the body into named poses (*asanas*), often involving stretching, strength, and balance, and is often provided to groups of people at yoga studios, led by a single expert practitioner. Some studios also incorporate chants and a spiritual focus to the practice, publicizing traditional Indian ideas about the body and health. Yoga attendees attend classes for a variety of reasons, ranging from appreciation for the body benefits of the stretching routines to a desire for fuller immersion in Indian belief systems, sometimes including visits to India.

Healthful practices are one type of complementary medicine, and the other broad category is consumable products. In the United States, practices would include the aforementioned acupuncture and yoga, in addition to massage, meditation, movement therapies, and relaxation techniques, among others. Consumable products include herbs and supplements, often for consumption but also for skin or other applications. Given all these types of available complementary medicine, the following section investigates the prevalence of these treatments among U.S. Americans, and some current trends.

Prevalence of Complementary Medicine in the United States

It is difficult to determine how many U.S. residents use complementary and alternative medicine, since there are differences in how studies use and define the terms. Researchers may ask people about holistic medicine or use the phrases unorthodox, unconventional, or marginal, which may be more or less inclusive of different treatments. For example, considerations of Navajo ceremonies typically incorporate the use of prayer as part of the treatment. In the case of Mexican curing ceremonies, should prayer also be included or is this something separate from medical treatment? These issues make it difficult to compare different studies and to consider rates of change in complementary medicine uptake or regional differences.

Notwithstanding, there are some surveys that have used terminology consistently and which provide several years of administration; these are particularly helpful for examining trends. A good example of a useful U.S. survey is the National Health Interview Survey (NHIS), organized by the National Center for Health Statistics, included in the Centers for Disease Control and Prevention. The study was initiated by the U.S. National Health Survey Act of 1956, and the first cross-sectional household-based survey occurred in 1957, with yearly iterations, although in 2019, the format was changed. Until 2018, it was based on questions formulated in 1997, and in 2002, 2007, and 2012, a subset of questions asked about complementary medicine. In 2007, the questionnaire was supplemented by interviews about

complementary medicine use by adults, included their reasons for usage.

The NHIS results for 2002 show complementary and alternative health usage by approximately 36% of U.S. American adults. Five years later, 38% of adults and 12% of children were accessing these treatments, and adults were spending US\$33.9 billion to fund products and practices. The most popular practice was deep breathing, with approximately 13% of the sample following this trend, followed by meditation, chiropractic, osteopathic treatments, massage, and finally, yoga, completing the more supported therapies, with approximately 6% of adults following this regimen.

In the same year, approximately 18% of these adults consumed herbal supplements. A 2012 estimate of the size of the nutritional supplements market worldwide suggests US\$32 billion in earnings, with extensive use of social media to reach consumers. As a result, many items that were once purchased only by intensive body-building competitors and marathon runners are now found on the shelves of typical households, and even children regularly consume energy drinks and bars.

The increase in family usage of complementary medicine correlates with a decrease in child vaccination levels. This effect has been identified in various studies, such as articles in the *American Journal of Public Health*, *Pediatrics*, and *Public Health Reports*. Each of these studies cites additional research documenting similar trends. As a result, different medical and health institutions try to deal with the low vaccination levels in various ways.

Many schools do not allow student registration unless vaccination schedules are completed, although increasing numbers of parents apply for vaccine exemptions, and nonvaccination rates are increasing, as shown in online links in Hill and colleagues' article in the *Morbidity and Mortality Weekly Report* and other links on the Centers for Disease Control and Prevention website. This trend is also reflected in universities recording many students without vaccinations in campus dormitories. These decreasing levels of vaccination compliance are accompanied by the spread of childhood diseases that were considered to be almost eradicated in the United States only a few years ago.

Several additional patterns have been identified in the use of complementary medicine. Generally, people who access conventional medical services at higher levels also tend to access alternative treatments with greater frequency. It also seems likely that people who use both types of service will attempt to source both when suffering a medical crisis and will then utilize whichever first becomes available. However, sometimes, people will turn more frequently to treatment with lower scientific support if they face too many barriers in sourcing standard medical services. These barriers may range from negative experiences with hospital outpatient procedures to denial of insurance coverage or other events, which may sour their views of conventional U.S. American medical provision.

Research also demonstrates that complementary medicine is applied in new and unusual ways, although it is difficult to document these trends unless the right questions are asked. For example, a 2008 study conducted by Jon C. Tilburt and colleagues, published in the *BMJ*, demonstrated that approximately half of 679 sampled internists and rheumatologists prescribed treatments that might be helpful for patients, although lacking in scientific proof of benefits. The analysis concluded that in these situations of long-term degenerative illness, physicians hoped to harness patient expectations in a placebo-like response. This unanticipated outcome in how doctors consider and apply complementary medicine hints that other unexpected results may occur in the future, given increasing efforts to bring complementary medicine into the fold of mainstream medicine through clinical tests of practice and treatment outcomes.

Contemporary Application of the Scientific Method to Complementary Medicine

Given the long-term usage of traditional medical practices and remedies and the increasing popularity of this style of medical provision in the developed world, it makes sense to evaluate the treatments, with the goal of diminishing the use of harmful practices and increasing understanding of beneficial ones. Perhaps, the best way of achieving this goal is to fund scientific studies, with controlled variables, random placement of research

participants into treatment and placebo groups, and accurate recording and publishing of results. There are many challenges in incorporating this strategy worldwide, but among the possible starting places may be the World Health Organization providing support for incorporating some aspects of traditional medicine into mainstream medical interventions. This offers an important way to supplement the inadequate number of trained medical doctors in many regions of the world and also a beginning in better documentation of local treatments utilized, with the aim of improving both mainstream health provision and complementary services.

In the United States, the U.S. Department of Health and Human Services established in 1998 a new National Institute of Health unit, the recently renamed National Center for Complementary and Integrative Health (NCCIH). As of fiscal year 2020, this unit controls approximately US\$151.9 million in funding and has supported numerous scientific investigations of products and practices with clearly documented results disseminated through academic publications and available to the general public online. The end result is improved public safety.

Some of the publicized research focuses on increasing awareness of beneficial practices for certain medical conditions. For example, in 2009, a study appeared in the academic periodical *Spine*, and also a simplified summary was placed online for general consumption. The investigation documented the value of Iyengar yoga for people with low back pain, decreasing pain levels, and improving body function and mood. In this manner, studies move complementary medical practices into the mainstream medical arena.

Other NCCIH-supported research focuses on diminishing harm, such as warnings about toxic materials in supplements. For example, a study conducted in 2005 and reported in 2008 in the *Journal of the American Medical Association* documented high levels of lead, mercury, and arsenic in a sample of Ayurvedic medicine purchased over the internet from Indian and American suppliers. Approximately one in five of the 193 tested products contained toxic metals, including those sourced from businesses extolling the benefits of their product testing. This type of information is extremely useful for those who order these products online.

In the future, it is likely that continued scientific testing of complementary medicine will result in better medical provision not only in North America but elsewhere, as well as a more efficient medical system, satisfying people from a variety of cultural backgrounds, with different conceptions of health and illness. This process will only be improved by increasing the numbers of minority graduates from mainstream medical schools.

Susan J. Wurtzburg

See also Belief Revision; Medicine, Contemporary; Worldview

Further Readings

- Alvord, L. A., & Cohen Van Pelt, E. (1999). *The scalpel and the silver bear: The first Navajo woman surgeon combines western medicine and traditional healing*. Bantam.
- Bleser, W. K., Elewonibi, B. R., Miranda, P. Y., & BeLue, R. (2016). Complementary and alternative medicine and influenza vaccine uptake in US children. *Pediatrics*, 138(5), e20154664. <https://doi.org/10.1542/peds.2015-4664>
- Hill, H. A., Elam-Evans, L. D., Yankey, D., Singleton, J. A., & Kang, Y. (2018). Vaccination coverage among children aged 19–35 months—United States, 2017. *Morbidity and Mortality Weekly Report*, 67(40), 1123–1128. <https://www.cdc.gov/mmwr/volumes/67/wr/mm6740a4.htm>
- Jesson, L. E., & Tovino, S. A. (2010). *Complementary and alternative medicine and the law*. Carolina Academic Press.
- Nebelkopf, E., & Phillips, M. (Eds.). (2004). *Healing and mental health for Native Americans: Speaking in red*. AltaMira Press.
- Perez, M. A., Raffy, R., & Luquis, R. R. (Eds.). (2014). *Cultural competence in health education and health promotion, 2nd ed.* Jossey-Bass.
- Saper, R. B., Phillips, R. S., Sehgal, A., Khouri, N., Davis, R. B., Paquin, J., Thuppil, V., & Kales, S. N. (2008). Lead, mercury, and arsenic in US- and Indian-manufactured ayurvedic medicines sold via the internet. *Journal of the American Medical Association*, 300(8), 915–923. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2755247/>
- Smith, P. J., Humiston, S. G., Marcuse, E. K., Zhao, Z., Dorell, C. G., Cynthia, H. C., & Hibbs, B. (2011). Parental delay or refusal of vaccine doses, childhood vaccination coverage at 24 months of age, and the

- health belief model. *Public Health Reports*, 126 (Suppl 2), 135–146. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3113438/>
- Spector, R. E. (2013). *Cultural diversity in health and illness*, 8th ed. Pearson.
- Tilburt, J. C., Emanuel, E. J., Kaptchuk, T. J., Curlin, F. A., & Miller, F. G. (2008). Prescribing “placebo treatments”: Results of national survey of US internists and rheumatologists. *BMJ*, 337, a1938. <https://www.bmj.com/content/337/bmj.a1938>
- Wang, E., Clymer, J., Davis-Hayes, C. C., & Buttenheim, A. (2014). Nonmedical exemptions from school immunization requirements: A systematic review. *American Journal of Public Health*, 104(11), e62–e83. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4202987/>

COMPLETENESS

Completeness is a property that theories can have when treated using the tools of modern mathematical logic and its mathematically precise treatment of logical consequence. This entry discusses two senses in which theories may be complete. Given the idealized view of theories logic employs, completeness in one sense has only been studied in relation to mathematical theories. But there is also a sense of completeness that applies to a theory of logical consequence, it is the completeness of a deductive system of logic.

Theories are a form of knowledge in which core concepts and assumptions are systematized so as to organize the understanding of the theory’s subject matter. The composition of theories in science can be very complicated. They may consist of data, experiments, models, and even research programs; they are not expressed in a uniform way. However, in logic theories are treated in an idealized and uniform way: All theories consist only of statements. Insofar as we can paraphrase the core elements of a theory into statements, this idealization is not misguided. Thinking of theories in these terms can help organize and reduce the number of core elements of a subject, since many other facts about the subject matter will follow directly from those core concepts. Logic studies a particular way that consequences can be drawn from statements, namely logical consequence.

Statements are related by logical consequence when one statement follows from a set of statements with certainty. For example, the Law of Gravitation predicts the velocity at which a stone hits the ground when it falls from a cliff. This is an elementary instance of one statement about a particular stone’s velocity following, with certainty, from a general statement together with other statements about the particular height of the cliff.

Background on Logic and Consequence

There are two distinct notions of completeness: one that applies to theories, and one that applies to a deductive system of logic. In order to understand either of these ideas, the concepts of a logical language and logical consequence must be clear.

A logical language is a recursively defined set of sentences that are built up from atomic components using logical connectives. Logical languages may be defined so that they allow the use of particular connectives or names for particular objects, and the introduction of particular predicates or multi-place relations. These are languages that can easily be adapted to describe applications of logic to the empirical world. The logical language of sentential logic (SL), also called *propositional logic* or *truth-functional logic*, uses simple sentence atoms as basic units and introduces special symbols for the words *and*, *not*, and others. First-order or predicate logic uses names, variables, and multi-place relations to form basic sentence atoms and adds to the symbols from SL symbols for *all* and *some*. Some languages might adopt a particular predicate, for example, $x < y$, as well as a particular range for its variables, for example, the positive integers.

The logical language defines sentences syntactically, and they may have a formal or an informal interpretation or meaning. The relation of *logical consequence* is a relation between sentences of a logical language, some defined as the premises and others as the conclusion. The premises of an argument may deductively prove one or more conclusions. The consequences of the premises are what *follows from* the premises.

As a result of the foundational work of Gottlob Frege in 1893, modern logic thinks of logical

consequence in two possible ways. The first is logical entailment, and the other is deductive proof.

Entailment

An *entailment* is a semantic (a truth and/or meaning based) relationship between a set of premises and a conclusion. Different semantic systems may be considered in philosophical or mathematical logic, but the classical relationship (which makes the common assumption that every sentence is either true or false and not both) of entailment is that the truth of the premises guarantees the truth of the conclusion. This is the standard definition of deductive validity. A distinct intuitionistic notion of entailment would result from eliminating the assumption that every sentence is true or false, construing true as “demonstrated or proved” and false as “refuted or disproven.” Entailment is understood in relation to assumptions about truth, and in relation to interpretations that make atomic sentences true. Classical logic uses classical interpretations, and intuitionistic logic uses distinct interpretations. To make things more general, let us represent the premises by a set of sentences, which we indicate with the symbol Γ and the conclusion by the letter φ . We then introduce the meaning of “ Γ entails φ ” in the following way:

$$\Gamma \models \varphi$$

if and only if every interpretation that makes all the members of Γ true, also makes φ true.

It would be possible to add subscripts to $\models$ if we needed to distinguish between different senses of entailment.

Deductive Proof

Logical consequence can also be understood as a relation determined by proofs. Proofs given in natural language, like those in mathematical research, are arguments which follow a set of rules and (may) use axioms to represent some fundamental facts about the subject matter of the argument. Proofs in formal logic are intended to carry out that project in such a way that the set of rules used in constructing the proof rely wholly on the symbols used in representing the premises and conclusion of the argument. In this way, formal

proofs are meant to capture the idea that the proof relation of an argument is based only on the form of the sentences (their syntax), and the logical words used, and not on their meaning or content. Such a system of proof avoids semantic intuitions about the meanings of the symbols and keeps unstated premises or *intuitions* from entering into the proof.

Relative to a language L , there are many methods that can be used to show that a set of sentences Γ can prove a sentence φ . These are so-called proof theories or deductive systems. Some systems represent proofs in a series of sentences each of which is either an axiom of the proof theory—a sentence which is taken for granted—or is derived from previous sentences in the series by rules which operate on their syntax. These are known as *Hilbert-style proof theories*. Other deductive systems do not use axioms; those systems have rules for the introduction of assumptions, and other rules which operate on the symbols of sentences derived from those assumptions or allow one to eliminate assumptions. These are known as *natural deduction systems* since they are meant to mimic the structure of proving things in natural mathematical contexts. There are other systems like so-called sequent systems or proof trees, but all of these create a rule-governed syntactic method to determine when φ is proved from a (possibly empty) set of sentences, Γ .

Regardless of its kind, once one has a system of deductive proof, LS , operating on a formal language L , it gives rise to a syntactic relation of logical consequence. We say that:

$$\Gamma \vdash_{LS} \varphi$$

if and only if there exists a proof of φ from Γ according to the rules of LS .

Theory Completeness

In mathematical logic, once one has fixed a logical language L , the theories in that logic are sets of sentences of L that contains all of their own logical consequences. In the terminology from above,

Γ is a *theory* in L relative to a notion of logical consequence $\models$

if and only if φ is a member of Γ when $\Gamma \models \varphi$.

When logicians speak of the theory of Peano arithmetic, for example, they mean the axioms of Peano arithmetic taken together with all of the logical consequences of those axioms. This notion can apply to theories in general, if we understand theories as collections of statements.

A theory Γ is *complete* when for each sentence φ of L , either $\Gamma \models \varphi$ or $\Gamma \models \neg\varphi$.

That means for each sentence φ either φ is entailed by the theory or φ is refuted by the theory. It is important to note that this notion of completeness is restricted to L , so although a complete theory will pass judgment definitively on each sentence, the content of those sentences will be restricted to whatever L can express. Complete theories tend to be fairly limited in complexity.

The theory of Presburger arithmetic is a complete theory, but it is an arithmetic that only deals with addition on the nonnegative integers: it cannot represent multiplication. And if multiplication were added to this theory, that could start to cause problems. One of the major results in logic is Gödel's Incompleteness Theorem of 1931. What Gödel proved was that if PA (Peano arithmetic qua theory) is consistent, then it is incomplete. So, provided PA is consistent, there are sentences φ in the language of PA such that neither $PA \models \varphi$ nor $PA \models \neg\varphi$. And PA is very nearly Presburger arithmetic with multiplication added.

The notion of the decidability of a theory or a logical system is different from the completeness of a theory or a logical system. A theory is decidable if and only if there is an algorithm that is usable in a finite time to determine whether an arbitrary sentence of L is a member of a theory or not. Whereas the completeness of a theory is about what sentences are entailed by the theory, the theory is decidable when given an arbitrary sentence, φ one can determine whether φ is entailed or refuted by the theory.

Completeness of a Logical System

The other sense of completeness is one that applies to logical systems. Recall the idea of deductive proof from above. A logical system's proof theory is complete when the proof theory relation $\vdash$ and

its entailment relation $\models$ are extensionally equivalent. That is, for all of the pairs of sentences or formulas Γ, φ in L ,

$$\Gamma \vdash \varphi \text{ if and only if } \Gamma \models \varphi.$$

When logicians prove completeness theorems or logicians say that a logic is complete, this is what is meant. A version of this result for first-order predicate logics was first proved by Kurt Gödel in 1930; that theorem is often referred to as *Gödel's Completeness Theorem*.

Strictly speaking, however, calling a logical system *complete* is an abuse of terminology. The completeness of a logical system is really about the proof theory exhausting the entailment relation of the system: if $\Gamma \models \varphi$ then $\Gamma \vdash \varphi$. The proof theory will not miss out on any of the valid arguments that the entailment relation says are valid. The other direction, if $\Gamma \vdash \varphi$ then $\Gamma \models \varphi$ is called the *soundness* of the proof theory. That means the proof theory will not lead one astray; all of the conclusions one can draw from Γ using the proof theory are entailed by Γ . It is very important to note that these notions of completeness only apply *relative* to a particular theory of consequence. It is the completeness of the proof theory with respect to a particular entailment relation or semantics for the language L . The completeness of a logical system says nothing, on its own, about the system's adequacy as a theory of valid argument.

There are many theories of entailment which reject central assumptions of classical predicate logic. For example, some logical theories relax the assumption that sentences may only be true or false. Some theories allow sentences to take on other truth values (many-valued logics) or even allow there to be sentences which are both true and false (paraconsistent logics). These semantic theories give rise to substantially different entailment relations over the same language as classical predicate logic. What is interesting is the proof theory for classical predicate logic is not incomplete with respect to these other entailment relations. The proof theory tends to be *unsound* with respect to those entailment relations. What the logician then does is devise new proof theories which are sound for the new notion of entailment/semantics.

Judy Pelham and Gillman Payette

See also Axiomatic theory; Deduction; Gödel's Incompleteness Theorems; Logical Theory

Further Readings

- Burns, S., Hägg, D., & Qian Z. (2022). *Sets, logic, computation: An open introduction to metalogic*. Retrieved April 18, 2022, from <https://builds.openlogicproject.org/>
- Gödel, K. (1986). *Collected Works I: Publications 1929–1936*. Oxford University Press.
- Hodges, W. (1993). *Model theory*. Cambridge University Press.
- Magus, P.D., & Button T. (2021). *Forall X: An introduction to formal logic, Calgary Remix*. <https://forallx.openlogicproject.org/forallxyyc.pdf>

COMPLEX SYSTEMS

Complex systems theory, sometimes referred to as *complexity theory*, does not originate from a single discipline and hence does not have a singular definition. In order to approach a useful theory of complex systems, we can identify shared definitions and typically attributed qualities in order to attain an overview, if not quite a consensus. This entry provides a working definition of a complex system, describes the structure of systems generally, and goes on to identify and explain the core concepts inherent in the study of complex systems and the theories that have been advanced by scholars in the various disciplines from which this field emerged. The entry concludes with a detailed discussion of the most widely recognized theoretical influences and movements related to and coming out of the study of complex systems.

A complex system is made of multiple components, which interact with each other over time, resulting in patterns of behavior at the system level and at other scales that cannot be predicted through a simple causal understanding of the individual parts and their functions. Furthermore, the organization of components and the relationships of parts to each other at individual and collective scales in terms of flows and interactions are as important as the components themselves for cognition of system behavior. In other words, typically reductive

approaches that are used to break down “complicated” problems into smaller parts, which are resolved separately, are unlikely to produce results in the study of complex systems or to unveil traceable causal chains resulting in collective dynamics.

Although this is an extremely generalized description, it will be expanded in order to demonstrate how this statement incorporates possibilities of emergent behaviors, self-organization, unpredictability, feedback loops, and sensitivity to initial conditions and in some cases outcomes based on learning, memory, and interaction with environments, similar to evolution. There is a clear distinction between “complex” systems and “complicated” systems. A number of extremely complicated systems exist, such as airplanes and large infrastructure projects such as dams. The aim and characteristics of these complicated systems are that they are predictable and linear: The same outcome is always expected for the same input, with any deviations from this treated as errors to be corrected. Complex systems, on the other hand, are typically nonlinear, and while it may be possible to identify the patterns of behavior, the system is not predictable in all aspects.

Many real-world systems demonstrate some or all behaviors attributed to complex systems. Examples include ant colonies, the biosphere and ecosystem, the immune system, social or political groups, the global economy, and cities as both artifacts and living adaptive systems. Theories from complex systems have gained popularity over the past few decades, partially due to the failure of existing theories to explain or predict real-world phenomena such as the global financial crisis of 2007–2008. They are also providing useful approaches to function and learn across different disciplines, such as physics, biology, economics, ecology, mathematics, computation, and economics, through identification of similar or overlapping phenomena. Complex systems theory emphasizes the impossibility of a single perspective or language to describe all the qualities of the system or a single entity to control all aspects of the system. It creates the need to reconsider the logic of reductionism, the absolutes of positivism, and the causality of determinism. While frames of reference and subjectivity from relativism are part of complex systems, lack of validity intrinsic to absolute relativism is not accepted as the default outcome. Instead,

systems of relations between objects, as acknowledged in G. F. W. Leibniz's relationist description of space and time and Joseph Kaipayil's reference to relational particulars, appear to provide a more useful operating framework.

The System in Complex Systems

A system can be described as a structure comprised of components or a group of elements that are interactive and/or interdependent and hence related to each other in some way. The components of a system have also been identified as entities, parts, or members depending on the context of use. Some descriptions of systems are based on the maintenance of functionality toward a goal, but in the context of complex systems, a singular or specific "goal" can be contested and interpreted. Examples of physical systems typically refer to assemblages of things and parts making up a whole, such as a railway system. If things are replaced with facts or principles, then the assemblage can be described as forming a system of thought or knowledge. A coordinated collection of steps or methods may form a system of organization. The definition of a system typically necessitates the possibility of differentiating it from other systems, even while acknowledging that systems can influence, overlap with, be contained in, and contain other systems. A currency such as the Euro is an identifiable system that is subject to the rules of international monetary systems and influenced in terms of value by trading between nations in foreign exchange markets. System boundaries can be defined by the system observer.

Systems typically demonstrate a degree of robustness. The idea of robustness is contestable due to disciplinary interpretations, but in a general sense, a system that can still be identified using the initial criteria in the face of changing external conditions can be considered temporarily robust. The same system may undergo the equivalent of a phase transition in thermodynamics to a different system or state when it is no longer robust. The discussion of robustness has parallels with the distinction Crawford Stanley (Buzz) Holling makes between engineering resilience and ecological resilience, where the former is the ability of a system to return to its initial state following external shocks and the latter is the ability of

a system to transform into a new state while still maintaining desirable functional attributes. In this context, the historic memory visible in many complex adaptive systems (CASs) can also be related to the idea of robustness as it implies that certain aspects of order are carried into new states.

Systems have inputs and outputs and are often divided into two categories, namely, closed systems and open systems. *Closed systems* are defined as systems that do not interact with their environment, thus remaining uninfluenced by externalities. *Open systems* have boundaries but exchange matter, energy, or information with their environment. The environment is often the collective term used for external and other systems, with a relationship to the defined system in consideration. In the real world, pure closed systems are assumed unrealistic and may only apply in abstract experiments.

Systems are also sometimes categorized as hard or soft systems. Hard systems (HSs) deal with well-defined problems in a systematic way. HS thinking treats quantifiable situations as problems that can be solved. Decision theory, forecasting, and simulation are all techniques used in this context—with individual actors or agents within the defined system not assigned complex motivations—and the outcomes aimed toward rational outcomes or at least known outputs. HSs thus deal with resolution. Soft systems methodology (SSM) starts with the assumption that the problem itself is not easily defined and hence difficult to quantify. SSM is used to address situations where people, motivations, and relationships are involved, and recognizes that the systemic understanding is dependent on the observer's objectives for the system. SSM thus addresses systemic cognition. Many real complex systems such as cities are made up of systems of systems and overlap the subjectivity of soft systems with the quantification of HSs, simply explained as the physical artifacts of a city being the outcome of human motivations.

Systems are defined or made apparent by the identification of their boundaries. Using systems theory, it becomes possible to connect components to components and systems to systems, and this can result in impenetrable definitions of extreme complexity that become impossible to analyze or utilize in any meaningful way. Hence, it is

important to define the boundary of a system with reference to context and function, with adequate awareness of environmental inputs and outputs. A system is typically identified by selecting the known entities that are inside the system and those that are outside, based on the ability to differentiate this identified group of entities from the others in terms of difference in type, structure, or collective behavior, and so forth.

Systems can also be understood as transformation processes with boundaries. This refers to a process or processes where inputs and outputs can be identified. While this last is often thought of as system used in mathematics or physics, where it may apply to thermodynamics, it can also be understood in terms of processes of immigration, in which the status of people's citizenship may change.

It is possible to study most if not all real-world phenomena as systems, through definition of these as linked systems, subsystems, and meta-systems. Complex systems are a special category of systems allowing the study of how relationships and processes between components result in unpredictable collective behaviors and how the system and its environment influence each other. Complex systems also enable the study of complex social and natural phenomena in mathematical and computational terms, bringing previously too-complex phenomena into the realm of scientific analysis.

Core Concepts: Nonlinearity, Emergence, and Self-Organization

Nonlinearity in systems is most strictly defined using two mathematical qualities. A system is formally nonlinear if it does not verify the properties of either superimposition or homogeneity. Both principles refer to the predictability of outputs based on known inputs. Hence, the principles remain unverified when output values do not behave as expected. The majority of real-world systems appear to be nonlinear, displaying qualities of positive and negative feedbacks, interference, and limitations in range when described using linear descriptions. Nonlinear systems are used in various disciplinary contexts, such as nonlinear regression in statistics, nonlinear optics in physics, and nonlinear population studies in ecology.

Emergence is the process by which smaller entities, components, and patterns interact with each other and in open systems with the environment, resulting in (typically larger) entities or behaviors not observable at the initial scale. This process is hence considered irreducible. In his definition, Jeffrey Goldstein, an economist, refers to the formation of novel and coherent structures and patterns during self-organization. Emergence can be seen as “strong” or “weak”; it is possible to simulate weak emergence using computers, but not strong emergence. Because identification of structure and pattern can be subjective depending on the definition of a system or choice of system scale, as can the interpretations of a pattern, emergence is often not considered to have a purely objective character. There appears to be general agreement, however, that emergence is related to synergetics—in other words, small-scale interactions resulting in a greater whole—and as in Hermann Haken's exploration of self-organizing structures in open systems, to multiple interacting subsystems. The theory of evolution provides a useful example and was used by George Henry Lewes, a philosopher of science, in his distinction between “resultants” and “emergents.” Resultants, according to Lewes, are phenomena that can be predicted by the preceding conditions. Emergents, on the other hand, as seen in evolution, refer to the possibility of completely new forms at various stages, despite being related to previous stages. There is a tendency to assume that emergence occurs in systems without top-down control, but certain degrees of organization, differentiation, and connectivity appear to be requirements for the process. Large numbers of interacting components—due to the higher number of interactions—are also often assumed to result in greater possibilities of emergence, but this may simply result in “noise,” which does not eventually result in emergent behavior and does not affect behavior at other scales. While there may be a minimum threshold of entities in specific systems for emergent processes to take place, it is the organization of the entities and types of relationships with particular reference to hierarchies and subsystems that appear to result in emergent outcomes. Emergent processes can occur due to both causal relationships and scalar interconnectivity. Phenomena observed in ant colonies and

flocks of flying geese are often used as the examples of emergence in action. One of the qualities of emergent systems is that they can interact with the environment as open systems and hence not follow the second law of thermodynamics; that is, thermal equilibrium through entropy.

Self-organization is closely related to emergence, but for the purposes of clarity, a distinction can be made by suggesting that emergence is a process more related to scale, while self-organization is more closely related to time. Self-organization is a process by which seemingly random structures and patterns can be seen to form in multientity systems without any observable or intentional top-down direction or controlling agent. The process can be observed in systems from many fields, such as physics, cognitive psychology, economics, sociology, medicine, and ecology. Behaviors such as swarming in biological systems or stock market crashes in economic systems can be studied as self-organizational processes. Negative and positive feedback loops amplify and dampen the fluctuations within a system, potentially contributing to the self-organized formation of structures and patterns. Although some definitions describe self-organization in the context of an isolated or uninfluenced system, biologists and ecologists acknowledge the relationship of organisms with their environment even during the study of self-organizational processes.

Complex Adaptive Systems

CASs are a particular group of complex systems, demonstrating some additional phenomena of particular importance to complexity in real-world systems. CASs—such as artificial intelligence (AI)—can exhibit the ability to learn from experience and evolutionary traits. Other qualities of CASs are sometimes identified as self-similarity, the ability of both agents in the system to adapt to changing conditions and the adaptability of the system itself. John H. Holland, a professor of psychology, computer science, and engineering, defined the term in a paper with the same name in 1992. He suggested that CASs not only have the ability to learn but may also demonstrate the ability to anticipate in the process of adapting to future conditions, thus adding yet another layer of

complexity. Many ecologies and human social group-based endeavors in a cultural and social system such as the global economy, political parties, or communities, including online tagging or social systems, can be studied as CASs. CASs can incorporate some or all qualities of complex systems (although sometimes named differently in disciplinary contexts, such as communication, cooperation, and spatial or temporal organization) and can be used in conjunction with game theory. Cities are increasingly used as examples of CASs. They are combined spatiotemporal and behavioral structures affected by, and affecting, the various individual and collective agents made up of citizens, socioeconomic subsystems, and sociotechnical endeavors, with multiple internal and environmental factors acting to influence the system at any one time, a system that is constantly in a state of becoming without a specific end state or aim.

Historic and Current Themes in Complex Systems

Complex systems, or the complexity sciences, have emerged over time through the identifiable overlapping influences of several disciplines and theories and remain an evolving body of knowledge today. Instead of attempting to identify specific historical stages leading up to the present state of complex systems, the following sections discuss, in roughly chronological order, the most widely recognized theoretical influences and movements related to and coming out the study of complex systems.

Chaos Theory

Chaos theory is the study of dynamic systems displaying seemingly random behaviors. Theoretically, chaotic systems are deterministic in that the same experiment should produce exactly the same results if the initial conditions are the same. The problem with this in practice is that absolute accuracy over an infinite length of time can only be theoretically achieved. In the 1880s, Henri Poincaré, a French mathematician also known as a theoretical physicist, engineer, and philosopher of science, experimented with a three-body system, a study of three or more interacting bodies that

attempts to determine the motions of three bodies over time, based on the knowledge of the initial masses, positions, and velocities of the three bodies. He found that the tiniest differences in initial conditions could result in dramatically different outcomes. Poincaré proved mathematically that increasing the accuracy of initial conditions when repeating the experiment did not result in proportional reduction in variation in the outcomes, which could continue to be dramatically different from the first experiment no matter how small the error, with the uncertainty in forecast increasing exponentially over time. In the 1950s and 1960s, Edward Norton Lorenz, an American mathematician and meteorologist, started to question the linear statistical models used in meteorology and the realistic accuracy of long-term predictions. While running computer models used to predict weather patterns, Lorenz chose to rerun a partially completed sequence and used the numbers noted from the end of the incomplete sequence, instead of starting from the beginning. He found that the results of his computer model were very different from expectations and deduced that this was because he had not used precisely the same numbers from the incomplete sequence but had instead rounded the numbers up or down by a decimal point. This led Lorenz to experiment with chaotic systems and conclude that the precision required to repeat predictable outcomes in such dynamic systems was possibly beyond the realm of human accuracy, resulting in the suggestion that prediction of past or present phenomena is in fact impossible. This theory was outlined in a 1963 paper for the New York Academy of Sciences and later presented at the 1972 meeting of the American Association for the Advancement of Science in Washington as what has since become widely known as the “butterfly effect.” Lorenz asked whether the flapping of a butterfly’s wings in Brazil could potentially result in a tornado in Texas. Although chaos theory is typically described in terms of extreme sensitivity to initial conditions, for a dynamic system to be termed chaotic, it also needs to be topologically mixed and have dense periodic orbits. Chaotic systems are unpredictable and yet deterministic in that the initial conditions determine future behavior. This means that despite their surface similarities, chaotic

systems and complex systems are usually differentiated.

Systems Theory

Systems theory, like complex systems, is not restricted to a single definition. It was developed in various disciplines, partly to allow for transdisciplinary frameworks and comparisons. Natural and human-related phenomena are the outcome of overlapping domains, and a systems theory approach attempts to provide a field of inquiry that allows this type of study and its reconciliation with traditional HSs. Systems theory may be considered a branch of systems thinking that enables the interdisciplinary consideration of areas previously treated separately by introducing the possibility of incorporating natural evolution and human subjectivity into the scientific study of all systems. As early as 1912 to 1917, Alexander Bogdanov, a Russian physician and philosopher, used the word *tektology* to refer to an approach unifying several disciplines from the social and biological to the hard sciences by understanding phenomena in terms of systems of relationships and organizational principles. Ludwig Von Bertalanffy (1901–1972) is widely accepted as having helped initiate the modern term, with his advance of general system theory, which he called *Allgemeine Systemlehre* and proposed as a solution to the fragmentation and duplication of research he had observed in various disciplinary fields. Both Bertalanffy, a philosopher by training, and Paul Weiss (1898–1989), a biologist, found that the existing mechanistic concepts inadequate for use with the natural systems that were part of their overlapping work. Bertalanffy clearly distinguished between *closed systems*—typically mechanistic systems—and *open systems*, which applied to living things by proposing that the laws of thermodynamics did not always apply to the latter. Early contributions to systems theory can also be traced to English mathematician and philosopher Alfred North Whitehead’s (1861–1947) view of the world as a network of interrelated processes, Talcott Parsons’s (1902–1979) action theory, and Niklas Luhmann’s (1927–1998) social systems theory. Systems theory influenced the development of ecological systems thinking, social systems theory, and systems science engineering.

Cybernetics

The French physicist and mathematician André-Marie Ampère first used the word *cybernetique* in 1834 in reference to civil government. However, it was Norbert Wiener (1894–1964) who popularized the word in its contemporary sense by writing a book called *Cybernetics: Or Control and Communication in the Animal and the Machine* in 1948. A Greek word meaning “steersman” provides the etymological basis for the term. Over time, multiple definitions for the field of cybernetics have come about due to its transdisciplinary appeal and usage. In a general sense, cybernetics is often referred to as a field concerned with the dual theories of communication and control, especially in the context of automatic control systems or self-correcting systems. However, a deeper study of cybernetics and its multiple definitions also allows inquiry into the directions the field has taken. Gordon Pask (1998–1996), an educational theorist, chose to interpret cybernetics as the art of manipulating defensible metaphors. Warren McCulloch (1898–1969), a philosopher, was interested in the internal and external communication of an observer within an environment. Gregory Bateson (1904–1980), an anthropologist, observed that cybernetics was focused on form and pattern. Stafford Beer (1926–2002), a management consultant, was interested in aspects of effective organization.

Although the multiple definitions for cybernetics expand the possibilities for application, most take into account the idea of how a system processes information and reacts to this information, with a chosen action or change to itself in order to improve either one or both of these tasks in the next iteration. Cybernetic systems typically refer to controlled systems, in which the organizational theorist Arthur J. Kuhn refers to detector, selector, and effector functions of a system. The detector is defined as concerned with the communication between systems, the selector is based on the rules of the system in terms of how decisions are made, and the effector concerns how transactions between systems take place. Examples of systems displaying cybernetic properties can be biological, cognitive, mechanical, social, and so forth. The nervous system is a biological example, while the autopilot system in an airplane is a mechanical example. Typically, cybernetics is associated with efficiency and identified

performative goals by being concerned more with what something does than what it is. This approach is the reason cybernetics is extremely transdisciplinary and potentially forms a meta-disciplinary language.

Some of the traditions that have existed in cybernetics can be compared side by side. Circular causality has been explored in the context of technology, in the form of regulation and control, and is particularly important in computation. Epistemology—the study of the nature and extent of knowledge—uses cybernetic ideas to study identity and autonomy in human social systems. Cybernetic systems are studied both in terms of how we observe them and in terms of how the systems observe themselves as well as other systems.

Systems theory and cybernetics are related historically and sometimes used interchangeably, but the origins of cybernetics are more associated with engineering fields, while general systems theory is thought to have emerged from a biological perspective. Cybernetics is associated with a wide range of phenomena and theories in the study of complex systems, such as cellular automata, game theory, and AI, the last of which emerged from cybernetics and later became distinct.

Artificial Intelligence

AI is a term coined by John McCarthy in 1955 as part of a summer workshop proposal at Dartmouth College. He defines it as the science and engineering involved in the making of intelligent machines. This definition has been generalized to the study and design of intelligent agents, where the term *intelligent agent* refers to a system that has the ability to perceive its environment and take action as a response. Whereas the original intention was to simulate human intelligence, this aspect known as “strong AI” has proved to be difficult in practice and remains a longer term goal. The two other aims, known as applied AI and cognitive simulation, have seen significant advances. Applied AI—also known as advanced information processing—is the field of development related to “smart” systems for commercial application. Cognitive simulation refers to the testing of theories about the functioning of the human mind, such as facial recognition.

The first widely acknowledged AI program was a checkers (draughts) program written by Christopher Strachey in 1951. This effort was advanced by Arthur Samuel, who by 1955 had developed a checkers program that could learn from experience, thus providing one of the first examples of evolutionary computing. John Holland's work in automating evolutionary computing has resulted in the advent of genetic algorithms.

Early work in AI is attributed to Alan Mathison Turing (1912–1954), who in his description of the “Turing machine” in 1935 provided the blueprint for the modern computer as we know it. Turing is also famous for formulating the Turing test in 1950, which is based on a computer convincing at least one third of designated interrogators into believing they are interacting with a human being. In 2014, a computer named “Eugene Goostman” passed the Turing test by convincing 10 of the 30 judges at the Royal Society in London. The validity of this test in proving a computer has displayed “human intelligence,” however, remains contested. In 1997, Deep Blue, a computer built by IBM, was able to defeat the world's reigning chess champion Gary Kasparov in a six-game match. The fact that human intervention was allowed to help Deep Blue in between games once again results in a debate about a computer's ability to learn, as opposed to simply processing huge amounts of information extremely quickly.

The primary areas of AI research include learning, problem-solving, reasoning, language, and perception. Research can be understood as two potentially distinct methods, the first being the “symbolic” approach and the second the “connectionist” approach. These can be thought of as top-down and bottom-up, respectively. While the symbolic approach has seen advances over the years, connectionist approaches have experienced renewed interest since the 1980s. The approach of training artificial neural networks embodied in the bottom-up approach has been applied in the context of visual perception, financial analysis, telecommunications, medicine, and language processing and is related to certain types of machine learning. Nouvelle AI, which describes systems referring to the external world instead of to an internal model only, and its related “situated”—embodied intelligences in the real world—approach, has refocused

attention on Turing's discussions on machine sensing from external environments.

Dynamical Systems

The theory of dynamical systems is relevant to the study of long-term system behavior using mathematical equations, particularly in mechanical or deterministic systems. Mathematical models are utilized to predict systemic change in time or dimension. Although the classic example is the predictability of the position of a clock in time, dynamical systems have also been used in the context of natural phenomena such as the changing number of fish in a lake over the seasons. Dynamical systems study today mainly focuses on chaotic systems and has been used over time to try and understand the movement of the solar system, weather, growth of crystals, traffic jams, and the motion of a billiard ball.

Agent-Based Modeling

Agent-based models (ABMs) are a specific type of computational model used to simulate and study agent behavior and the resulting effect on a system. Agents in this context can be individual or collective entities and typically interact within an environment they can perceive. The basic idea is that simple behavioral rules result in complex behaviors. ABMs are widely used in the social sciences and contain elements of AI, computational sociology, multiagent programming, emergence, and game theory. Possibilities of probability and randomness can be introduced into agent interactions.

Fuzzy Logic

Fuzzy logic is applied when the categorization of things into distinct sets cannot be accomplished easily. The term was introduced by Lotfi A. Zadeh in 1965. A typical example is an attempt to define the exact point when a statement changes from false to true, from black to white within a smooth gradient. Fuzzy logic states that there is an infinite gray area between black and white, which is defined partially on the basis of subjectivity and context. People are typically able to assign a value

to the truthfulness of a statement using numbers between 0 and 1 or the modifiers “quite true” or “mostly false.” Zadeh suggested that precise statements lost their meaning as complexity rose, thus positing that fuzzy categorization was more appropriate for complex systems.

Data Mining

Data mining is the process of analyzing data from multiple perspectives and either summarizing it into understandable information structures or discovering useful patterns and correlations. While the term is often used to refer to large-scale information processing, the useful applications overlap with machine learning and its supervised and unsupervised learning methods. Machine learning is a branch of AI, which has come to the forefront with the advent of Big Data. Examples of machine learning can be found in search engines’ ranking websites and junk mail filters learning usage. Medical databases are currently being created and analyzed to create knowledge for diagnosis and cures, and recent developments in computer vision and natural-language processing have begun to demonstrate direct translation of real-world knowledge into digitally recognizable data without the need for tagging or prior categorization.

Case-Based Reasoning

Case-based reasoning (CBR), or case-based modeling, attempts to find solutions to new problems based on the solutions of similar problems in the past. CBR can be observed in many fields, such as the use of legal precedents in law, and is becoming a popular methodology for both computer reasoning and human problem-solving. It is useful to observe the innate learning ability of the method as new cases are solved and additional cases are incrementally added to the systems solutions data. This last quality relates CBR to CASSs.

Ulysses Sengupta

See also Artificial Intelligence; Cybernetics, 20th Century; Developmental Systems Theory; Ecology;

Evolutionary Psychology; Game Theory; Physics, Quantum Theory; Systems Science; Systems Theory

Further Readings

- Alexander, C., Ishikawa, S., & Silverstein, M. (1977). *A pattern language: Towns, buildings, construction*. Oxford University Press.
- Allen, P. M. (1997). *Cities and regions as self-organizing systems: Models of complexity*. Taylor & Francis.
- Bateson, G. (1970). Form, substance and difference. In Favareau, D. (Ed.), *Essential readings in biosemiotics* (pp. 501–518). Springer.
- Batty, M., & Marshall, S. (2012). The origins of complexity theory in cities and planning. In E. Stolk, E. Tan, H. Meyer, & J. Portugali (Eds.), *Complexity theories of cities have come of age* (pp. 21–45). Springer.
- Beer, S. (1966). *Diagnosing the system for organisations*. John Wiley & Sons.
- Checkland, P. (1999). *Systems thinking, systems practice: Includes a 30-year retrospective*. John Wiley.
- Cilliers, P. (1998). *Complexity and postmodernism: Understanding complex systems*. Psychology Press.
- de Roo, G., & Silva, E. A. (2010). *A planner's encounter with complexity*. Ashgate Publishing.
- Érdi, P. (2007). *Complexity explained*. Springer.
- Goldberg, D. E., & Holland, J. H. (1998). Genetic algorithms and machine learning. *Machine Learning*, 3(2), 95–99.
- Goldstein, J. (1999). Emergence as a construct: History and issues. *Emergence*, 1(1), 49–72.
- Haken, H. (1977). *Synergetics*. Springer.
- Holland, J. H. (1992). Complex adaptive systems. *Daedalus*, 121(1), 17–30.
- Holland, J. H. (1999). *Emergence: From chaos to order*. Da Capo Press.
- Holling, C. S. (1996). Engineering resilience versus ecological resilience. In Schulze, P. (Ed.), *Foundations of ecological resilience* (pp. 51–66). National Academies Press.
- Kaipayil, J. (2009). *Relationalism: A theory of being*. JIP Publications.
- Kauffman, S., & Macready, W. (1995). Technological evolution and adaptive organizations. *Complexity*, 1(2), 26–43.
- Kuhn, A. J. (1986). *Organizational cybernetics and business policy: System design for performance control*. Pennsylvania State University Press.

- Lewes, G. H. (1877). *Problems of life and mind*. Trübner & Company.
- Lorenz, E. N. (1963). Deterministic nonperiodic flow. *Journal of the Atmospheric Sciences*, 20(2), 130–141.
- Luhmann, N. (1982). The world society as a social system. *International Journal of General Systems*, 8(3), 131–138.
- McCarthy, J., Minsky M. L., Rochester, N., & Shannon, C. E. (2006). A Proposal for the dartmouth summer research project on artificial intelligence, August 31, 1955. *AI Magazine*, 27(4), 12.
- McCulloch, W. (1965). *Embodiments of mind*. MIT Press.
- Naess, A. (1973). The shallow and the deep, long-range ecological movement. *Inquiry*, 16, 95–100.
- Parsons, T., Shils, A. E., & Smelser, N. J. (1965). *Toward a general theory of action: Theoretical foundations for the social sciences*. Transaction Publishers.
- Pask, G. (1961). *An approach to cybernetics* (Vol. 8). Harper & Row.
- Prigogine, I. (1967). *Introduction to thermodynamics of irreversible processes*. (3rd ed.), Interscience.
- Von, B. L. (1950). An outline of general system theory. *British Journal for the Philosophy of Science*, 1, 134–165. <https://doi.org/10.1093/bjps/I.2.134>
- Von, B. L. (1968). *General systems theory* (Vol. 38). Braziller.
- Whitehead, A. N. (2010). *Process and Reality*. Simon and Schuster.
- Wiener, N. (1961). *Cybernetics, or control and communication in the animal and the machine*. (2nd ed.). MIT Press.
- Wolfram, S. (1984). Universality and complexity in cellular automata. *Physica D: Nonlinear phenomena*, 10(1), 1–35.
- Wolfram, S. (1998). Cellular automata as models of complexity. In L. Lam (Ed.), *Nonlinear physics for beginners: Fractals, chaos, solitons, pattern formation, cellular automata, complex systems* (pp. 197–202). World Scientific Publishing. (Original work published 1984 in *Nature* 311, 419–424)
- Zadeh, L. A. (1965). Fuzzy sets. *Information and Control*, 8(3), 338–353.

COMPUTER SCIENCE

See Big Data; Cybernetics, 20th Century; Software Engineering

CONCEPT

A concept is akin to a general or abstract idea. The philosopher Hilary Putnam famously said that concepts are signs used in a certain way. This is similar to the 18th-century philosopher David Hume's notion that an abstract idea is an exemplar along with a rule of use. Whereas those thoughts are powerful, they can be hard to grasp at first, and we can think of concepts in a variety of ways. A concept can be an intuited object of thought or it might be directly conceived. When someone says that they know a concept, then it means that they comes across one *in action*, because it becomes larger and greater than its label as a word, phrase, clause, or sentence. Generally, every idea, object, issue, process, person, and place can develop into a concept. A standard example is the concept of triangle. The concept of triangle is not, itself, a triangle but instead allows us to reason about all triangles by reasoning about it. We can learn about all triangles by investigating the properties they all share, and we do that by investigating the concept. It is an integral part of human cognition, and various approaches and viewpoints are presented by the scholars to define it. Concepts cannot appear in isolation as they are employed to represent knowledge. During any such practical representation, concepts remain associated with other.

Definitions of Concepts

The term *concept* was derived from the Latin word *conceptum* (*that which is conceived*). The explanation of a concept has been dealt with in the subject matter of psychology, linguistics, and philosophy. Many theories, sometimes fundamentally contradicting each other, have been proposed about the nature of concepts over time. When researchers try to define explicitly what a concept is, they may fail insofar as a universally accepted definition of a concept does not exist. However, it is generally recognized that concepts are somehow associated with the process of human cognition. Moreover, it is commonly accepted that concepts are universal and abstract. Since concepts do not enclose the differences between or specifics of the objects to which they apply, they are considered

abstract. Therefore, all objects to which a concept applies, known as *extensions* of concepts, are treated as indistinguishable from the respective concept. The universality of concepts arises from the fact that they apply to every object in their extension. The abstract nature of concepts does not affect the nature of objects in the extension of a concept. These objects can thus be real or imaginary, atomic or composed, abstract or concrete. A concept can be anything existing in the human mind, such as a strategy or way of thinking, or shared through behavior and human language, including actions and tasks.

Concepts in Philosophy

The ancient Greek philosopher Aristotle was the first to provide the classical theory of concepts. Aristotle's *theory of definition* defined a concept by comparing a pair of concepts which he termed as *genus* (a sort, kind, or family) and *differentia* (a distinguishing characteristic). His theory defines the concept *human*, for example, as a *thinking animal*. The *animal* determines the family human belongs to, which corresponds to genus, and it is *thinking* that distinguishes humans from other animals, and thus corresponds to differentia.

Philosophers Arthur Schopenhauer and John Locke consider individual experience, in the form of sensation and reflection, to be responsible for the formation of concepts. They considered concepts to be abstractions. The concepts abstracted from human experience are also known as *posteriori* concepts. Another type of concept introduced by the philosopher Immanuel Kant. He mentions pure or *a priori* categories or concepts that originate in the human mind. The concepts of time and space are Kant's examples of *a priori* concepts.

The philosopher, logician, and mathematician Gottlob Frege argued that we comprehend the world around us through concepts. He provides illustrations of a symbol (*sign*), an object (*reference*), and a concept (*sense*) to show the differences between them. The symbols *morning star* and *evening star* refer to the same object: the planet Venus. However, the concepts of these two phrases are entirely different. The former can be observed in the morning, while the latter is visible in the evening. The two signs represent two different senses which correspond to different observations of the

world, namely the time and space of observation in the *star* example.

Similarly, John Sowa considers a concept as a mediator that relates a sign to its reference. This mediation can be presented with the triangle, popularly known as a *semiotic triangle*, put forward in the 1920s by I. A. Richards and Charles K. Ogden. If we want to show a person and her activity in the triangle, the lower two corners show an icon resembling a person and the other corner will show the printed symbol representing her name, for example, Mary. The peak of the triangle will represent a *neural excitation*, for example, how someone thinks of her at a particular time and space. This *neural excitation* is a mediator between the symbol and its object, and it is referred to as a *concept*. In the field of philosophy, concepts are directly associated with the perception of the world and essence of human existence. But in other fields of science, this treatment is more pragmatic. Concepts are expressed in terms of the field's terminology, depending upon the way concepts are employed in practice.

Concepts and Other Scientific Disciplines

In language science, a concept is generally considered as a unit of meaning, often formed through the abstraction of concrete phrases and words that correspond to the conceptual meaning. For example, WordNet, a lexical database, illustrates the linguistic treatment of concepts. This database interconnects nouns, verbs, attributes, and adverbs in the English language. For concepts, WordNet defines them through sets of synonym words called *synsets*. A synset consists of synonym words, such as man, human being, homo, and human, which define a concept that can be lexically expressed by any of the words in the synset. Additionally, WordNet provides an explicit definition in the description of the synset *concept*. The WordNet example shows that a concept can be represented with more than one lexicon. Moreover, when individual words are represented with more than one meaning, then they can stand for more than one concept. On the other hand, the significance between words and concepts can be understood by the fact that concepts are language-independent, unlike words. When doing translation between different languages, the semantic

aspect of words is expressed and preserved to a large extent. Beside several other definitions of concepts in linguistics, some define it as a common meaningful shared entity in the mind of users.

In engineering, a concept is concerned with building models of entities from reality. Such perspective defines concepts as creatures of the computational field, which presents through their representations in software such as XML and UML and in systems of axioms. In defining concept, mathematics, unlike engineering, transcends the reality that is recognized through our intuition and emotions. Mathematician Carl Benjamin Boyer defines mathematical concepts, such as *derivative* and *integral*, as abstract mental constructs. He states that such concepts are beyond the world of sensory experience; however, they may be observed by abstraction from nature or by intuition.

Concepts and Knowledge Representations

As discussed in the previous section, concepts are often considered as atomic elements in knowledge representations. For that reason, it is worth knowing how concepts are treated in this domain. Certain formalisms represent organized knowledge as concept maps and define concepts as perceived regularities in objects or events, or records of objects or events, termed as *labels*. For most concepts, the label is a symbol or a word. One formalism is description logic, used for representing logic-based knowledge through classes (concepts), relations (roles), and objects (individuals). The concepts denote sets of individual objects in description logic.

Frames are knowledge representation structures influenced by the organization of human memory. Contrary to other many formalisms for representing knowledge, concepts are not regarded as atomic units when represented by frames. They are considered as sets of highly structured entities which can be marked with recursive structures comprising pairs of slots (also known as *attributes* or *properties*) and their values. Object-oriented languages are closely related to frames. They discern two elemental structures: classes and their instances, known as *objects*. A class specifies the methods and properties that can be applied to

control the properties of all the objects that are instances of a special class. The set of methods and properties in a class allow for those objects possessing a behavior and a state. As classes act as abstractions of concrete objects, they stand for concepts, whereas objects represent instantiations of concepts.

Concepts and Implicit Definitions

It is important for a machine—as in the instance of the lexical database discussed earlier in this entry—to recognize and define concepts consistently and exactly. If such a machine fails, for example, to recognize which text represents concepts and which does not, then it becomes difficult to make concepts recognizable so that they can be extracted automatically from any type of texts.

A solution is proposed by a vector space model (VSM). In a VSM, each document corresponds to a vector with coordinates representing the frequency of the detected index items in a document. The number of dimensions of each vector corresponds to the number of index terms detected in a document collection. The measurement of two documents, such as the level of similarity between two documents, can be done by calculating the inverse function of the angle between them or the inner product of the corresponding index vectors when the similarity function is at a maximum and the angle between two vectors is zero. The measurement of documents can also be observed in terms of similarity when a new index item is given to a document collection. If there is a decrease in similarity level, then the newly assigned index item has a high discrimination value, and it is considered as a *good* index term. The opposite of this is true for a *bad* index term. For this reason, the *good* index terms can be known as *concepts*, as they represent even the smallest units of knowledge bearing as much meaning as possible; in other words, this representation is enough to decrease the similarity level between the documents when given to a document collection.

The field of word-sense disambiguation, especially automatic identification of word senses (also known as word-sense disambiguation), also provides more implicit definitions of concepts. The word clustering approach to the word-sense induction involves clustering of semantically

similar words, such as they can produce a specific meaning and can thus be considered as concepts. The similarity between different words can be determined by the observation of the syntactic dependencies of specific words in actual texts, for example, words functioning as grammatical subjects of a transitive verb, direct objects of a ditransitive verb, adjectives, and so on. In this way, concepts can be defined implicitly as a group of words that share a high degree of syntactic dependencies. Many other approaches and algorithms exist. The commonality between them is the clustering or grouping of words showing distinct meanings. These clusters are known as *concepts*. The exactness of words in defining a concept, however, varies, and it depends upon the definition of similarity that a particular method uses for grouping words.

As we have seen, there are various possible ways through which we can define concepts. In the following section, we will discuss how concepts can be organized and thus be made predictably available for use.

Concepts and Their Relationships

It is necessary for concepts to be linked with each other, and they cannot be isolated for any practical use in representing knowledge. They are the chief means of organizing concepts. The relationships that link one concept to another are employed in day-to-day communication and the significance of identifying the underlying relationships between concepts can be understood with the following example. The statement “Mary has an Intelligence Quotient of 140” clearly states a very explicit relationship that Mary has some characteristic named Intelligence Quotient that is equal to 140. Nonetheless, the statement implies many other implicit relationships. These relationships include things like the abbreviation of Intelligence Quotient as IQ or the value 141 follows 140 and it is preceded by 139, or Mary is commonly a female first name of a human, or that an IQ equal to 140 shows a person of high intelligence, and so forth. But if context is lost, then many things described in this statement are also lost. We will not get any meaning in isolation, for meaning is the sum total of relationships. Here, semantic relations are important. They are the basic means of organizing

concepts into knowledge structures. The meaningful associations between two or more entities or concepts are also known as *semantic relations*.

Valence is a basic property of semantic relations. It includes the number of concepts a semantic relation can associate. The valence of a semantic relation is expressed with the number of fields, slots, places, or sides of the relation. Mostly, the relations connect two concepts, and they remain binary. However, the relation of *send* is, for example, a ternary relation, connecting the one who sends, the one who receives, and that which is sent. Relations where there is higher valence than two can be decomposed into more than one binary relation. As mentioned earlier, all relations can be represented as concepts that are linked with a single *link* relation.

The pioneering linguist and semiotician Ferdinand de Saussure distinguishes two types of basic semantic relations: syntagmatic and paradigmatic relations. A *syntagmatic* relation is a relation between two neighboring concepts in the texts, such as a syntagmatic relation is the relation “loves” linking the concepts “John” and “Mary” in the sentence “John loves Mary.” This relationship holds for an ad hoc association of concepts in an actual text. By contrast, a *paradigmatic* relation is a relation between concepts that is independent of the actual use of the respective concepts in the texts, such as the concepts *loves* is inherently associated with the concept *act of loving* by the paradigmatic relation *is an act of*. Some formalisms for concept organization recognize a special type of relation known as an *attribute* for Semantic Web ontology languages and object-oriented modeling languages. The properties of concepts and their extensions are denoted by attributes. They are generally represented as a feature of entity that is employed to extension and model concepts. The attributes can be classified into four groups: Certain values remain attached and associated with the concept. *Class attributes* assign such values to the concept, and thus remain the same for all instances of a concept; *instance attributes* assign different values to each extension of a concept; *local attributes* share same-name with attributes that are attached to different concepts; and *global attributes* are applied to all concepts, for example in ontology, in a specific conceptual structure.

Concepts and Graphical Organizations

Semantic relations often result in a graph-like structure while organizing concepts. Such graphic structures are commonly described with a formalism, for instance, concept maps, semantic networks, and conceptual graphs. A concept map is comprised of concepts and the relationships between them. The relationship is expressed with phrases and words mentioned on the corners, linking concepts. This relationship forms meaningful statements, or propositions about constructed or naturally occurring events or objects in the world. A conceptual graph has two modes: conceptual relations and concepts. Any graph can present concepts in a number of ways such as rectangles and circles. The conceptual relation links with a concept at the edge of the conceptual graph. The organization of concepts remains hierarchical in concept maps. The general concepts remain on the top, whereas specific concepts are located in the bottom.

Amitabh Vikram Dwivedi

See also Abstract Knowledge; Category Theory; Conceptual Analysis; Conceptual Blending; Explanation; Generalization; Information Theory; Knowledge; Theories, Semantic Conception of

Further Readings

- Baader, F., & Nutt, W. (2002). Basic description logics. In F. Baader, D. L. McGuinness, D. Nardi, & P. F. Patel-Schneider (Eds.), *The description logic handbook* (pp. 47–100). Cambridge University Press.
- Berners-Lee, T., Hendler, J., & Lassila, O. (2001). The semantic web. *Scientific American*, 284(5), 34–43.
- Boyer, C. B. (1959). *The history of the calculus and its conceptual development*. Dover Publications.
- Cañas, A. J., & Novak, J. D. (2009). *What is a concept map?* <http://cmap.ihmc.us/docs/conceptmap.html>
- Ding, L., Finin, T., Joshi, A., Pan, R., Cost, R. S., Peng, Y., Reddivari, P., Doshi, V. C., & Sachs, J. (2004, November). Swoogle: A search and metadata engine for the semantic web. In *Proceedings of the thirteenth ACM international conference on information and knowledge management*, Washington, DC, United States.
- Fellbaum, C. (Ed.). (1998). *WordNet: An electronic lexical database*. MIT Press.
- Finin, T., Ding, L., Pan, R., Joshi, A., Kolari, P., Java, A., & Peng, Y. (2005). Swoogle: searching for knowledge on the semantic web. In *Proceedings of the national conference on artificial intelligence (AAAI)*, Pittsburgh, PA, United States.
- Frege, G. (1892). Über Sinn und Bedeutung. In *Zeitschrift für Philosophie und philosophische Kritik* [English edition: Frege, G. (1980). On sense and reference (M. Black, Trans.). In P. Geach & M. Black (Eds.), *Translations from the philosophical writings of Gottlob Frege*] (3rd ed.). Blackwell.
- Gomez-Perez, A., & Corcho, O. (2002). Ontology languages for the semantic web. *IEEE Intelligent Systems*, 17(1), 54–60.
- Halladay, S., & Milligan, C. (2004, January 5–8). The application of network science principles to knowledge simulation. In *Proceedings of the 37th annual Hawaii international conference on system sciences*, Big Island.
- Locke, J. (1975). *An essay concerning human understanding*. Oxford University Press. (Original work published 1690)
- Miller, G. A. (1995). WordNet: A lexical database for English. *Communications of the ACM*, 38(11), 39–41.
- Ogden, C. K., & Richards, I. A. (1923). *The meaning of meaning*. Harcourt, Brace & Co.
- Petersen, W. (2007). Representation of concepts as frames. In J. Skilters, F. Toccafondi, & G. Stemberger (Eds.), *Complex cognition and qualitative science. The Baltic international yearbook of cognition, logic and communication* (Vol. 2, pp. 151–170). University of Latvia Press.
- Princeton University (2010). *About WordNet*. <http://wordnet.princeton.edu>
- Salton, G. (1989). *Automatic text processing: The transformation, analysis, and retrieval of information by computer*. Addison-Wesley.
- Salton, G., & Wong, A. (1975). A vector space model for automatic indexing. *Communications of the ACM*, 18(11), 613–620.
- Saussure, F. de. (1916). *Cours de linguistique générale* [English edition: Saussure, F. de. (1983) *Course in General Linguistics* (R. Harris, Trans.)]. Open Court, La Salle.
- Schopenhauer, A. (1851). *Parerga and paralipomena* [English edition: Schopenhauer, A. (1974) *Parerga and paralipomena: Short philosophical essays* (E. F. J. Payne, Trans.)]. Oxford University Press.
- Smith, B. (2004, November 4–6). Beyond concepts: Ontology as reality representation. In A. C. Varzi & L. Vieu (Eds.), *Proceedings of the international*

- conference on formal ontology and information systems (pp. 73–84). Turin.
- Sowa, J. F. (1984). *Conceptual structures: Information processing in mind and machine*. Addison-Wesley.
- Sowa, J. F. (2000). Ontology, metadata, and semiotics. In *Conceptual structures: Logical, linguistic, and computational issues. Lecture notes in computer science* (Vol. 1867, pp. 55–81). Springer.

CONCEPTUAL ANALYSIS

Conceptual analysis, also known as *explication* among empirical researchers, is a fundamental logical step of scientific inquiry and theory development. It builds a detailed and comprehensive definition of a concept, normally also including empirical measurements or indicators of that concept. In science, particularly social science, no theory can be built without such a comprehensive clarification and articulation of the concepts that are needed to establish a hypothesis, which turns to be a theory if it passes substantial empirical tests. Nowadays, many academic debates are the result of unclear concepts.

Conceptual analysis is philosophically rooted in linguistic philosophy, which holds that language is systematically misleading and therefore needs to be carefully analyzed. With roots that can be traced to the beginnings of philosophy and logic, conceptual analysis was highly exalted by analytical philosophers such as Rudolf Carnap (1891–1970), who held that scientific concepts must be analyzed and defined by philosophers. In the 1950–1960s, however, philosophers like W. V. O. Quine (1908–2000) and Hilary Putnam (1926–2016) were convinced that analysis to produce a priori concepts was not possible nor necessary. Conceptual analysis has been undergoing a revival in philosophy in recent decades, widely attributed to Australian philosopher Frank Jackson's work. According to philosophers including Stephen Stich and Jonathan Weinberg, Jackson's 1998 book *From Metaphysics to Ethics: A Defense of Conceptual Analysis* provides perhaps the most sophisticated defense of the use of conceptual analysis in philosophy that has ever been offered. In it, Jackson argues that the very business of telling a story

in the vocabulary of daily life requires grasping the essential structures of the concepts and therefore the exercise of conceptual analysis, which is an important tool for understanding the material world and the foundation of categorization, meaning construction, and communication. All theory construction is dependent on conceptual analysis. In science and social science, researchers have also developed guidelines for conceptual analysis in all empirical studies.

Concepts, Constructs, and Variables

Concepts, in their widely adopted sense, are the basic unit of thoughts and statements. They are mental representations, categorizations, and labels for things being observed in the real world. Ontologically, concepts have been deemed by philosophers (e.g., John Locke, David Hume, Jerry Fodor, and others) as mental representations/images (lately popular among psychologists and other social scientists), abilities particular to cognitive agents (see, e.g., Michael, Dummett), or Fregean senses/abstract objects (see, e.g., Christopher Peacocke). The classical theory of concepts, which is held in the highest regard among all theories of concepts, maintains that all concepts are expressed in languages and their definitions consist of simpler concepts. Definitions of complex concepts beyond common sense are mostly acquired through paradigmatic conceptual analyses. In science, conceptual analyses are ongoing as a theory is developed.

An important task for social scientists in building social science theories is to test and identify causal relationships between or among concepts. In social science, those concepts are related to human reflection and the society. *Constructs*, a term popularized at least partly by education researcher Fred Kerlinger in his oft-cited book *Foundations of Behavioral Research*, are a special type of concepts invented by researchers for the purpose of scientific inquiry; concepts usually have multiple dimensions. For Kerlinger, *weight* is a concept, but *intelligence* is a construct. Other scholars, for instance, communication researcher Steven Chaffee, well known in part for his monograph *Explication*, simply use the word *concepts* in both situations that Kerlinger mentioned. The

monograph is in fact a part of the series *Communication Concepts*, which examines communication concepts that have long attracted scholarly interest. All concepts should be carefully analyzed before being used to build propositions (hypotheses). In *Explication*, Chaffee showed how the concept *age* needs meticulous examination. In some cases, ages reported in years and those calculated with birth dates are not even correlated.

Concepts, including Kerlinger's constructs, are not directly observable. To study the concepts, social scientists, particularly those who use statistical techniques in research, need to analyze and define them, and develop one or more indicators from them. Those observable and measurable indicators mostly contain variation and are therefore called *variables*. The lengthy and replicable process of searching for the definition of, and developing indicators from, concepts is called *explication*.

The Process of Explication

Several philosophers have talked about explication, such as Immanuel Kant and Edmund Husserl, but Rudolf Carnap is most well known as the one who valued explication in scientific theory construction. He defined explication as the logical analysis and construction to clarify a concept that is vague or not exact in daily life or an earlier stage of science. For old concepts, explication is conducted through searching out new definitions that are clearer, simpler, more exact, and, above all, fit into a systematic structure of scientific concepts. The more a concept is brought into the system of other scientific concepts through explication, the more fruitful and useful the concept is to construct theories or laws. Explication, Carnap said, is one of the most important tasks of philosophy.

Social scientists have divided explication into two steps: conceptualization and operationalization. Conceptualization, what communication researcher Steven Chaffee called *meaning analysis* in his *Explication*, is the process of articulating the essential meaning of the concepts. The meaning of a concept, according to Chaffee, can be defined in two ways: by distillation and by list (similar to the traditional logic terms *intension* or *connotation* and *extension* or *denotation*). *Distillation* consists

of the abstract statements of the essential elements of the concept (intension or connotation). For instance, media dependency is a media effect that when audience have no access to other communications on a topic they depend maximally on the mass media. In definition by list (extension or denotation), all concepts subordinate to the concept that needs definition are listed. The concept of mass media, Chaffee suggested, would be better to be defined by list as including newspapers, books, magazines, television, radio, and films. A distillation definition of the concept *mass media* is not easy to produce, as scholars are unclear whether the *mass* refers to mass production, mass audiences, or both.

Authors who use a concept without conceptualization may adopt different meanings of the same concept in the same literature and thus confuse themselves and readers. Chaffee and his coauthor found that some communication researchers used the term *social reality* to refer to the real social situation that existed even though people in the situation might not see it, while other communication researchers used the same term to refer to the false reality socially defined by people in it. The understanding of the relationship between communication and the *social reality* in this literature, therefore, cannot be solid.

Operationalization is the step, based on conceptualization, of finding empirical definitions (also called *operational definition* by some social scientists) for the concepts. Those empirical definitions can be directly observed and measured, and represent the essence of, but never the whole, concepts. Finding those empirical definitions is the ultimate goal of explication. No concept, according to empiricism, can be scientifically studied without an empirical definition that can be directly measured. In psychology and education research, studies on how to develop empirical definitions or measurements of concepts constitute an important and fundamental subfield. Obtaining adequate measurement methods, psychologist Jum Nunnally argued, is the major problem in the science of psychology. Nunnally believed that psychological theories were filled with concepts that could not be accurately measured, which hindered the development of psychological science. Communication researcher Sharon Dunwoody also mentioned that a concept

measured inaccurately in her research nearly invalidated a line of her work for 15 years and made her aware of the significance of explication in social science research. For a long time, she studied how the concept *uncertainty*, which she conceptualized as public perception of what the scientists didn't know about an issue, impacted the information use of members of the public for the purpose of making risk judgments about that issue. They measured that concept by asking respondents about their level of uncertainty, and later found that for most of the respondents the word *uncertainty* actually meant how much they themselves didn't know of the issue.

Many instances of confusion in social sciences have arisen from malfunctioning conceptualization and/or operationalization of concepts. The dispute between second-level agenda-setting researchers and framing-setting researchers in the mass communication field, for instance, is largely caused by unclear definitions of the concepts *agenda* and *frame*. *Agenda*, a good metaphor, is accurately operationalized as the number of stories in the original agenda-setting study conducted by Maxwell McCombs and Donald Shaw. But when studies citing agenda-setting theory accumulate, particularly when the agenda-setting theory is extended to its second level (from objects to attributes), the concept *agenda* becomes ambiguous and is defined in a variety of different ways. The same is true of *frames*. Although frames have been conceptualized as the way of defining meanings in the texts and operationalized as rhetorical devices, patterns of message organizations, word choices, tone, and so on, other scholars have observed that some framing researchers just operationalize frames as issue topics. As a result of these ineffective conceptualizations and operationalizations, the salience of a story is used in some studies as a measurement of second-level agendas but in other studies as a measurement of frames, thus leading to the dispute that each school claims that the other school is a part of its own research. In the field of international relations studies, David Baldwin found that much theoretical controversy had been generated due to the unclear explication of the concept *dependence* and related concepts such as *interdependence* and *dependency*. Marketing researcher Charles Young also finds that different

marketing research companies, such as Ameritest, MSW, Milward Brown, Ipsos-ASI, and ARS, use different concepts not only to document campaign effects but also measure the same concepts, such as attention, recall, motivation, persuasion, and brand linkage, in totally different ways. Those different measurements of the same concepts, Young notes, generate confusion in the industry.

Dimensions and Levels of Measurement

For empirical researchers, a successful explication of a concept should end with a measurable empirical definition, a variable that can be used in data collection and statistical analysis. A variable can have one or more than one observational point (indicator). What is more, some concepts, called constructs by some researchers, contain different perspectives (dimensions), and each perspective may demand several indicators. For instance, through literature review, Alice Eagly and Shelly Chaiken identified three dimensions in human attitudes: cognitive, affective, and behavioral, each of those perspectives having several indicators. Although single-item measurement is often seen in scientific research, to reduce measurement errors in order to close the gap between empirical definitions and the concept that they measure, multiple-item measurement for a single-dimension variable or for a dimension of a multiple-dimension variable is strongly suggested in social science.

Scientific investigators usually measure variables at four levels: nominal (categorical), ordinal, interval, and ratio. Nominal level is the lowest; each number only represents a category and the numbers cannot be compared for a ranking. A common example of a nominal-level measurement is gender. Male = 1 or female = 1 is just a labeling of different genders and has no numerical meanings. The second level, ordinal, enables comparison and ranking, but values at this level still cannot be mathematically calculated in any way, because values between data points are unknown. For instance, movie A, B, C, and D are ranked as 1, 2, 3, and 4 with 1 being the most favorable ranking. The measurement clearly shows the level of favorability of those four movies. But one cannot say that the difference of

favorability between A and B is exactly the same as that difference between C and D. Interval level, beyond ordinal level, makes addition and subtraction possible but not multiplication and division. Interval scales place greater emphasis on a number's relationship with other numbers than on the numerical values of the numbers. Examples of interval level of measurement include temperature and IQ score. John's IQ score of 108 is 8 units higher than David's IQ score of 100, and that difference is smaller than the difference between Josh's IQ score of 102 and Mike's IQ score of 98. If the whole IQ scale from 0 to 200 is systematically changed to 300 to 500 or 700 to 900, no information contained in the IQ measurements will be changed, because numbers in interval scales have only comparative values and no absolute numerical values. At the ratio level, such systematic change of a whole scale is no longer feasible. Age and income are the examples normally measured at the ratio level. Changing Hank's \$4 to \$9 and Bob's \$3 to \$8, even though the comparison value between Hank and Bob's money has not been changed, the absolute values of their moneys have. In attitudinal research, scholars tend to use Likert-type scaling, Guttman scaling, Thurstone scaling, and semantic differential scaling to measure concepts. Those scales normally contain five to nine measuring points. Some conservative scholars analyze data collected with those scales at the ordinal level, but most scholars use them as interval-level measurements.

Reliability and Validity

Because the explication process goes a long way from the unobservable conceptual world to the measurable real world, information may get lost during the journey of conceptual analysis and errors may occur. All empirical measurements inevitably have some deviation from the concepts that they represent. In psychology and other social sciences, the accuracy of those operational definitions as results of conceptual analyses is documented by two indicators: reliability and validity.

Reliability documents how consistently the measurement is perceived, consciously or unconsciously, as an operational definition of the concept being studied among people. It is therefore

also called *dependability*, *stability*, *consistency*, or *predictability* by some researchers. Reliability essentially examines how many errors are produced simply because of misunderstanding of the operational definition or the ways of using those operational definitions. A reliable operational definition will be interpreted the same way from time to time and situation to situation by different respondents, whereas an unreliable one will be interpreted differently at different times or in different situations or by different people. Reliability can be tested in several ways, such as testing the same operational definition for multiple times or analyzing the internal relationship of components of an operational definition. In most areas in social science, when multiple items are used to explain a concept, a commonly used coefficient of reliability is *Cronbach's alpha*, which measures the internal consistency among the items that serve as indicators of that concept.

Validity documents the deviation of the operational definition from the real, consensual meaning of the concept being defined. It checks whether or not the operational definitions really measure what the researchers think that they are measuring, the concept. Based on how the deviation is checked, measurements have different types of validities. A joint committee of the American Psychology Association, The American Educational Research Association, and the National Council on Measurement Used in Education classify validity into four categories: content, predictive, concurrent, and construct. If a group of experts examine the operational definition and judge it as an accurate and comprehensive representation of the concept, then the operational definition has content validity. Predictive validity and concurrent validity both check the operational definition against an outside criterion to see if the empirical definition represents the concept accurately and comprehensively. There are therefore both called *criterion validities* also. The only difference is the times of the formation of the criterion and the checking of the operational definition against the criterion. For predictive validity, the outside criterion will be formed in the future, and so will the comparison of the operational definition against the criteria. But for concurrent validity, the outside criterion exists in the current time

and can be used to check the validity of the operational definition immediately. For instance, to check if a test to identify characteristics of a great journalist really measures what it claims to measure, a sampled group of journalists can be asked to take the test. If journalists who have done well in their job score high on the test and those who have done poorly score low, the test has concurrent validity. If the test is administered to a sampled group of college journalism students, and years after graduation those who score high on the test perform well as journalists while those who score low perform poorly, then the test has predictive validity.

Construct validity is based on the examination of a concept with related concepts, sort of like Carnap's system of scientific concepts. Construct validation is possible when two conditions are met: (1) From literature or experiences researchers know the relationships of the concept being analyzed with other existing concepts; and (2) those related concepts have known valid operational definitions or measurements already. To check the construct validity of the operational definition of a concept, researchers examine the relationships of that operational definition with the known valid operational definitions of concepts related to that concept. If the relationships between that operational definition and the other operational definitions are the same as the known relationships between that concept and the related concepts, the construct validity of that operational definition for that concept is then established. Otherwise, the conceptual analysis for that concept needs to continue until a valid operational definition is established so that the concept can be further studied in a hypothesis or a theoretical system, namely, a model. Construct validation is, as Kerlinger said, like using the whole process of developing a theory to develop an empirical measurement for a concept. In research practice, Kerlinger recommended to use factor analysis to check the validity of the measurement of a concept or examine the correlations among the measurements of two or more concepts. Recently, most researchers prefer using confirmative factor analysis or *structural equation modeling*, also called *latent variable modeling*, *path model*, or *covariance structural modeling*, to conduct construct validity analyses.

For instance, in an article published in *Molecular Psychiatry* in 2010, K. S. Kendler and M. C. Neale used a structural equation model, which they called a *path model*, to analyze the relationship between the construct *endophenotype* and other related concepts. Other methods of construct validity analysis also exist. For example, a group of researchers from Bowling Green State University, Annette Mahoney, Kenneth Pargament, Nalini Tarakeshwar, and Aaron Swank, used a meta-analysis of 94 studies published in the 1980s and 1990s to clarify the linkage of the concept of religion with other concepts in the family domain.

Finally, an unreliable operational definition cannot be a valid one, but reliability also does not guarantee validity. A valid definition, however, theoretically should be a reliable one. Unreliable or invalid operational definitions of scientific concepts can damage any scientific investigation. Having valid definitions of its fundamental concepts marks the maturity of a scientific field. To achieve that maturity is still a heavy duty for emerging areas of social scientific research. Even in some areas considered well developed, such as media effects research, as James Potter said when he attempted to explicate the concept *media effects*, the continual clarification of the central concepts in those areas is still the essential scholarly task.

Qingjiang (Q. J.) Yao

See also Deduction; Induction

Further Readings

- Chaffee, S. H. (1991). *Explication*. Sage Publications.
- Cronbach, L., & Meehl, P. (1955). Construct validity in psychological test. *Psychological Bulletin*, LII, 52(4), 281–302.
- Dunwoody, S. (2005). Explicate, please. *MAPOR News*, Fall issue, p. 4.
- Eagly, A., & Chaiken, S. (1993). *The psychology of attitudes*. Harcourt Brace Jovanovich.
- Kerlinger, F. (1964). *Foundations of behavioral research*. Holt, Rinehart, and Winston.
- Nunnally, J. (1987). *Introduction to psychological measurement*. McGraw Hill.
- Young, C. (2008). *The advertising research handbook* (2nd ed.). Ad Essentials, LLC.

CONCEPTUAL BLENDING

Conceptual blending is the popular name for a theory of idea creation and emergence that draws heavily from the work of Arthur Koestler in the 1960s and is most popularly associated with the work of Gilles Fauconnier and Mark Turner. Attempting to describe the functional and theoretical aspects of how humans think creatively and develop ideas, conceptual blending has recently been adopted in various fields as a framework for developing software and artificial intelligence, educational approaches, scientific inquiry, and, perhaps most commonly, linguistics. Methods for visualizing complex information have also started to incorporate conceptual blending as a way of addressing metacognition and intention in learning processes.

Matrices and Bisociation

Working to address issues of human creativity, Arthur Koestler (1905–1983) published *The Act of Creation* in 1964. In his book, Koestler conceives of frames containing individual concepts, whether creative or factual. These frames, which Koestler refers to as *matrices*, can be thought of as Euclidean two-dimensional planes (imagine a sheet of paper) which bound and contain the concept entertained in the mind of a thinker. The ability to entertain these individual concepts is not necessarily a creative act. In Koestler's model, the creative process occurs when two or more of these planes intersect. The point(s) of intersection between the matrices are the locus of the creative act wherein two or more concepts interact with one another in the mind of the thinker. Koestler refers to this active process of conceptual interaction as *bisociation*:

I have coined the term 'bisociation' in order to make a distinction between the routine skills of thinking on a single 'plane' as it were, and the creative act, which [. . .] operates on more than one plane. The former can be called single-minded, the latter double-minded, transitory state of unstable equilibrium where the balance of both emotion and thought is disturbed. (p. 36)

The bisociation approach to considering conceptual interaction encourages the investigation of

topics like narrative, humor, art, and science through a lens of emergence rather than delineated cycles of start–finish. The interaction of matrices can itself produce new matrices and in doing so contributes to an evolutionary model of concept creation and adaptation. Koestler applies this view to scientific inquiry explaining the additive process of connecting one instance of research with another to create new findings, theories, and concepts:

In the discoveries of science, the bisociated matrices merge in a new synthesis, which in turn merges with others on a higher level of the hierarchy; it is a process of successive confluences towards unitary, universal laws. (p. 39)

Mental Spaces and Cross-Space Mapping

The contemporary model of conceptual blending is most notably attributed to the work of Gilles Fauconnier and Mark Turner. Their 2002 book *The Way We Think: Conceptual Blending and the Mind's Hidden Complexities* draws heavily from the concepts Koestler identifies. Their work is predicated primarily on linguistic constructions and uses a network model (connections between related concepts) to conceptualize how people entertain multiple concepts and synthesize new ones. In this model, Fauconnier and Turner conceive of mental spaces as “small conceptual packets constructed as we think and talk for purposes of local understanding” (2003, p. 58). These mental spaces are contextual in nature and often have short life spans.

When mental spaces merge, they can aid in constructing new conceptions passing through phases Fauconnier and Turner refer to as *cross-space* mapping. This process progresses to a final step they term as *blend*. This final step is important to understanding concepts but more important to the creation of new ones. In most cases, the term *conceptual blending* refers to the process of mental spaces colliding in these cross-space maps analogous to the intersecting matrices in Koestler's model.

For example, if one person is attempting to describe a platypus to another person, they might describe it as a combination of a duck's bill and a beaver's body. For the duration of the conversation (or even shorter), the individuals may construct mental spaces where they can consider each

of these animals. Cross-space mapping takes place when the individuals entertain the implications of these mental spaces interacting with one another. Thus, in considering a platypus as a duck/beaver hybrid, conceptual blending leads to a mental image in the minds of the individuals. This blending does not guarantee accuracy, the process is inherently subjective, but a strong understanding of the process can help to improve creative and abstract reasoning as well as an individual's ability to engage in more systematic processes, such as scientific research, by drawing connections between dissimilar ideas or concepts.

Fields of Application

Although in its infancy, computational methods of investigating conceptual blending are used in research about user-interface design in computer software and structuring artificial intelligence systems. Some research into humor and metaphor engages conceptual blending theory. In a paper published as a result of a conference, State of Research Into Critical Thinking, it was clear that educational communication required skills and approaches beyond classic delineations of communication or linguistics. Conceptual blending has begun to be used as the basis of pedagogical strategies in educational literature.

As the traditional line between print and digital culture blurs, conceptual blending becomes increasingly useful in the study of linguistic approaches to communication and information exchange. Various approaches to linguistics and cognitive processing have employed conceptual blending as a framework through which to address human thought processes. Seana Coulson discusses conceptual blending in her 2001 book *Semantic Leaps: Frame-Shifting and Conceptual Blending in Meaning Construction*, in which she connects the approach to theoretical models of scripting and framing, among others.

Conceptual blending also connects two important aspects of learning about concepts (and by extension theory): metacognition and intention. Metacognition refers to thinking about one's own thoughts and intentionality has to do with students who engage in willful steps to construct meaning (see Hennessey, 2003; also Brown 1978; Flavell, 1976; Sinatra & Pintrich, 2003). In order

to support these aspects of the learning process, it is important to occasionally step outside of the traditional linguistic constructs often used to consider human communication and inform the various traditions of scientific inquiry.

James Sosnoski (2015) uses conceptual blending as the basis for a systematic approach to discourse analysis aimed at improving modern communication strategies in educational contexts. He connects metacognition and intention while challenging linguistic approaches in higher education such as lectures, objective tests, and textbooks. Learners develop conceptual categories as a blend of new ideas and common conceptual frames from typified past experiences. In some cases, *visualizing conceptualizing* may provide opportunities for matrices or mental spaces to interact in ways difficult to communicate verbally: "The model of conceptual blending provides a way of structuring visualizations that mirror the cognitive process" (2013, p. 10).

Gordon S. Carlson

See also Abstraction; Concept; Framework; Mental Models; Theory Construction

Further Readings

- Brown, A. L. (1978). Knowing when, where, and how to think: A problem of metacognition. In R. Glaser (Ed.), *Advances in instructional psychology*. Erlbaum.
- Coulson, S. (2001). *Semantic leaps: Frame shifting and conceptual blending in meaning construction*. Cambridge University Press.
- Fauconnier, G., & Turner, M. (2002). *The way we think: Conceptual blending and the mind's hidden complexities*. Basic Books.
- Fauconnier, G., & Turner, M. (2003). Conceptual blending, form and meaning. *Recherches en Communication*, 19, 57–86.
- Flavell, J. H. (1976). Metacognitive aspects of problem solving. In L. B. Resnick (Ed.), *The nature of intelligence*. Erlbaum.
- Hennessey, M. G. (2003). Metacognitive aspects of students' reflective discourse: Implications for intentional conceptual change teaching and learning. In G. M. Sinatra & P. R. Pintrich (Eds.), *Intentional conceptual change*. Erlbaum.
- Hunt, E. B. (1962). *Concept learning: An information processing problem*. Wiley.
- Koestler, A. (1964). *The act of creation*. Penguin.

- Nosich, G. M. (1991). A model for the national assessment of higher order thinking. In R. Paul (Ed.), *Critical thinking: How to prepare students for a rapidly changing world*. The Foundation for Critical Thinking.
- Sinatra, G. M., & Pintrich, P. R. (2003). The role of intentions in conceptual learning. In G. M. Sinatra & P. R. Pintrich (Eds.), *Intentional conceptual change* (pp. 1–18). Erlbaum.
- Sosnoski, J. J. (2013). *Visualizing conceptualizing: Conceptual integration learning apps*. Society for Conceptual Logistics in Communication Research. https://www.academia.edu/309747/Visualizing_Conceptualizing_Conceptual_Integration_Learning_Apps
- Sosnoski, J. J. (2015). *Reading and analyzing discourses*. Society for Conceptual Logistics in Communication Research. <http://www.sclcr.com/documents/ReadingAndAnalyzingDiscourses.pdf>
- Turner, M. (1997). *The literary mind*. Oxford University Press.

CONSILIENCE

Consilience, the unity or *jumping together* of knowledge, is a term referring to the linking of facts, fact-based theory, and causal explanations across disciplines. The term was drawn from the phrase *consilience of inductions*, introduced by the Cambridge scholar William Whewell in 1840 to describe a scientific method by which the colligation of facts and convergence of various types of evidence results in the strongest scientific conclusions. Consilience did not reach prominence, however, until it was resurrected and expanded in the 20th century by evolutionary biologist Edward O. Wilson. Based on Charles Percival Snow's 1959 lecture explicating two divided academic cultures, the humanities and sciences, Wilson advocated for the erasure of the *territorial* lines between disciplines in favor of cooperative exploration; in other words, consilience, the unification of scientific and humanistic knowledge. Although Whewell's idea of the consilience of inductions sparked much debate among philosophers of science, Wilson's conceptualization of consilience has generated even more widespread attention and controversy.

Consilience as Scientific Method

In 1840, the English philosopher and historian of science William Whewell wrote a two-volume synthesis of science titled *The Philosophy of the Inductive Sciences, Founded Upon Their History*. In the second volume, Whewell introduced the phrase *consilience of inductions*, which he used to describe the phenomenon of altogether different classes of fact jumping together at the same point, resulting in the establishment of the *best* scientific theories. Whewell noted that the most monumental scientific discoveries were better explained by the colligation of distant facts and observations than by individual suppositions, even those constructed for the specific purpose of explaining the phenomenon. In following, Whewell argued then that the theories with the greatest explanatory power are consilient—the results of the greatest unification of knowledge.

Whewell acknowledged that the reality of science is that facts are not framed all at once; rather, theories are generally added to over time, with various advances and subsequent research. Consilient theories are those which subsequent suppositions serve to simplify and harmonize; ones that become even more coherent as they are extended. Theories which are not consilient become more complex and unmanageable as they are added to. Consilience, therefore, characterizes theories which represent a test of necessary *truth* and have a tendency toward simplicity and unity while maintaining deductive strength.

At its core, the consilience of inductions suggests a scientific method. Put simply, the consilience of inductions is the idea that multiple sources of evidence cohering to the same conclusion strengthen that conclusion. The strongest consilience is provided when the methods used to produce the evidence are completely or vastly different, as is generally the case when data are gathered by different disciplines with their own unique research techniques. In other words, the fewer characteristics shared by evidence producing the same conclusion, the more confident researchers can be of that conclusion, even when one source of evidence is less reliable than the others. The consilience of inductions as a method for achieving confidence is particularly important to

disciplines or research areas in which experimental research is not possible (e.g., evolutionary biology/psychology and intimate partner violence research). When experimental evidence is lacking, multiple sources, diverse evidence, and various research strategies converge to form *consilience arguments* used to support causal inferences drawn from nonexperimental data and to rule out competing explanations.

An example of this was presented by Kwok Leung and Fons J. R. van de Vijver in their discussion of epidemiological research. As they note, epidemiologists generally follow the methods of Whewell's consilience of inductions, primarily because experimental data are nearly impossible to gather, since assigning study participants randomly to treatment and control groups is unethical (i.e., you cannot introduce an epidemic for research purposes). As such, Leung and van de Vijver outlined the four criteria epidemiologists rely on to draw causal inferences from nonexperimental data: (1) rule out alternative explanations based on effect size (the strength of association between cause and effect); (2) establish temporal sequence—cause preceding effect—to rule out alternative explanations; (3) consistence, a consilience criterion referring to replicable effects across different populations and in different circumstances; and (4) coherence, another consilience criterion referring to the absence of conflicting evidence for proposed causal relationships.

Consilience of Academic Cultures

During the 1959 Rede Lecture at the University of Cambridge, C. P. Snow argued that intellectual life, particularly in the West, was divided into two opposing cultures, the sciences and the arts/humanities. He cautioned that such a division limits our collective ability to solve the world's problems. Snow later suggested the creation of a third culture, comprised of scholars across various disciplines who share an understanding of the sciences; in particular, biology, which as he suggested was the least exact of the hard sciences and thus open to the most interpretation.

In 1998, as a response to the division in academia outlined by Snow, Edward O. Wilson

published *Consilience: The Unity of Knowledge*. Wilson proposed consilience on a much broader and more ambitious scale than Whewell did, defining consilience as the unification of knowledge and the key to erasing the divide between the humanities and sciences. To Wilson, the pursuit of consilience represents the ultimate academic and intellectual challenge, the result of which will be an understanding of human nature with a higher degree of certainty than any discipline could produce on its own. Cooperation between the humanities and social sciences and equal contribution from all disciplines and specialties will increase the diversity and depth of knowledge, and give purpose to intellectuals (i.e., the pursuit of consilience). Wilson further addressed Snow's assertion that division limits the ability to solve the world's problems by framing consilience as the only way to confront the two foremost problems facing humanity in the 21st century: overpopulation and environmental destruction.

Wilson's conceptualization of consilience has received a great deal of criticism, particularly by scholars in the humanities who take issue with Wilson's assertion that establishing or refuting consilience is only achieved through methods developed in the natural sciences. As Wilson argued, the success of the natural sciences is in fact due to scientific reductionism. Opponents argue that scientific knowledge can be equally informed by humanistic study, and that methodological favoritism promotes inequity as opposed to a true alliance between the humanities and sciences. Fueling this argument is the fact that a large portion of the work toward consilience has been undertaken by scholars examining aspects of the humanities from scientific perspectives. An example of one of the few exceptions is the work of Jonah Lehrer, who contends that works of visual, musical, and culinary arts revealed truths about how the brain functions decades before these truths were obtained through reductionist scientific methods. To illustrate, Lehrer offers the early 20th-century novel *In Search of Lost Time* by the French novelist, critic, and essayist Marcel Proust (1871–1922). Lehrer argues that Proust constructed a model of the workings of memory seemingly consistent with the findings of recent scientific studies,

conducted over 100 years after the publication of Proust's novel.

Another area of controversy regarding consilience is the placement of the social sciences, which includes disciplines such as sociology, political science, psychology, and communication. Wilson acknowledged social sciences as a separate entity from the humanities and natural sciences but argued that the social sciences will divide themselves between the natural sciences (becoming continuous with biology, in particular) and the humanities, a process he states has already begun. Wilson does not suggest these discipline will disappear, only that they will exist in a radically different form. Evidence demonstrating Wilson's prediction can be found in many social science disciplines or subdisciplines. Examples include the convergence of sociology and biology (led by Wilson himself) into *sociobiology*, and the convergence of communication and biology into *communibiology*. Despite these convergences, there remain strong opponents who argue that the social sciences should represent a third distinct academic culture and not be relegated to the natural sciences or the humanities. Jerome Kagan presented such an argument against consilience and in favor of three academic cultures. Kagan outlined how the assumptions, terminology, and contributions of each of the three cultures are unique. As such, the contributions of each culture cannot be applied to any other culture, because the sources of evidence are so vastly different.

Michelle E. Pence

See also Hypothetico-Deductivism; Logic, Inductive; Theory Across Disciplines

Further Readings

- Kagan, J. (2009). *The three cultures: Natural sciences, social sciences, and the humanities in the 21st century*. Cambridge University Press.
- Lehrer, J. (2008). *Proust was a neuroscientist*. Mariner Books.
- Leung, K., & van de Vijver, F. J. R. (2008). Strategies for strengthening causal inferences in cross cultural research: The consilience approach. *International Journal of Cross Cultural Management*, 8(2), 145–169. <https://psycnet.apa.org/doi/10.1177/1470595808091788>
- Snow, C. P. (1963). *The two cultures and a second look: An expanded version of the two cultures and the scientific revolution*. New American Library.
- Whewell, W. (2012). *The philosophy of the inductive sciences: Founded upon their history* (Vols. 1 & 2). Cambridge University Press. (Originally published 1840)
- Wilson, E. O. (1998). *Consilience: The unity of knowledge*. Vintage Books.

COSMOLOGY

The domain of cosmology is nothing less than the universe (or cosmos) at large, meaning everything that has, has had, or will have an existence, whether space, time, matter, or energy. In this all-embracing sense, cosmology goes far back in time in the form of mythological and philosophical speculations. What is far from obvious is that the universe can also be described and understood by means of observations related to scientific theories, but this is a late recognition which only bore fruit in the 20th century, especially after physicist Albert Einstein in 1915 promulgated the fundamental field equations that still serve as the theoretical basis of the modern understanding of the universe. Cosmology as a recognized scientific discipline only dates from the 1960s, when the big bang theory, so-called, won broad acceptance, and since then the standard big bang model of the universe has developed into a highly advanced science supported by sophisticated observations and experiments. Nonetheless, cosmology in the early 21st century is still a science in progress, and the standard model is challenged by several alternative views of the universe. In what follows, the development of cosmology is traced from the formulations of Isaac Newton, which were based on the heliocentric system of Copernicus put forth in the 16th century. The entry continues with an account of the 19th-century advent of astrophysics following the invention of spectroscopy; the theory of thermodynamics and the dispute between the proponents of the steady-state and the entropic heat-death hypotheses of the universe; and in

the 20th century until the present day, the model of the expanding universe and its confirmation with the discovery of the nebular red shift, lending support to the *big bang* theory of the universe's origins. The entry concludes with a review of the current consensus and the enigmas posed by dark matter and dark energy.

The Newtonian Universe

During the first half of the 17th century, an increasing number of astronomers adopted the Copernican world system, formulated by the Polish astronomer and mathematician Nicolaus Copernicus (1473–1543), with the Sun situated in the center of the universe and the Earth degraded to one planet out of six. Far outside Saturn, thought to be the outermost planet, resided the multitude of stars, about which practically nothing was known. How far away were the stars and how many were there? Did they all have roughly the same distance from the Sun, or did they stretch indefinitely out in endless space? The new cosmological system that emerged in the early 18th century owed much to Isaac Newton's mechanical physics, as presented in his masterpiece *Principia* from 1687. Not only were Newton and his contemporaries able to estimate the distances to the nearest stars, but based on Newton's universal law of gravitation, they also formulated a new cosmological model which would survive for about two centuries.

According to this model or worldview, space was infinite, uniformly filled with stars, and governed by gravitational forces. However, Newton and a few other natural philosophers vaguely realized that even with an infinite number of stars the universe would slowly become unstable and eventually come to a catastrophic end by collapsing upon itself, a scenario they considered to be unacceptable. The so-called *gravitational paradox* was only investigated in detail at the end of the 19th century when the German astronomer Hugo von Seeliger proved that a homogeneous and infinite stellar universe could not be brought into agreement with Newton's law of gravitation. Still, Seeliger's conclusion was not generally considered a serious problem for the Newtonian worldview. One way to escape the gravitational collapse

problem was to introduce a slight change in Newton's law at very large distances and another was to keep the law intact but postulate a *hierarchical universe* with a particular inhomogeneous distribution of stars.

In any case, the great majority of astronomers disregarded cosmological questions of this kind, which they saw as speculative and scarcely scientific. In the spirit of positivism, they argued that only the observed part of the universe belonged to science proper and that this part was limited to the myriads of stars contained in the Milky Way galaxy, to which our solar system belongs. The American fin-de-siècle astronomer Simon Newcomb admitted that space might be infinite and even populated with an infinity of stars and nebulae, but he added that this was merely a philosophical hypothesis which the scientific astronomer could safely ignore. Cosmology in the grander sense of the term—the study of the universe in its totality—was widely looked upon with suspicion and cultivated by only a minority of astronomers and physicists. After all, the universe as a whole is not observable, and according to the norms at the time science was concerned solely with what can be observed.

Astrophysics and Thermodynamics

Before the mid-19th century, the term *astrophysics* was close to being an oxymoron. After all, how could one possibly know about the physical constitution of the stars, not to mention their chemical composition? The invention of spectroscopy, the central technique of the new astrophysics, would entail a profound transformation of astronomy and in the long run, of cosmology. It was known that the observed wavelength of a light signal increases (shifts toward the red end of the spectrum) with the relative velocity of the source, and by the 1880s astronomers using the spectroscope were able to detect the *red shift* of stars and in this way to determine their radial velocities. In the 1910s, the American astronomer Vesto M. Slipher discovered the analogous red shift effect for nebulae, a discovery that turned out to be of momentous importance to cosmology. However, the cosmological interpretation of the nebular red

shifts—that they indicate an expanding universe—only became clear at about 1930.

The theory of thermodynamics was based on two fundamental laws of which the first was energy conservation and the second was that the entropy of a closed system inevitably increases. According to the German physicist Rudolf Clausius, one of the founders of thermodynamics and the inventor of the concept of entropy, the two laws were valid for the universe as a whole. Since entropy is a measure of a system's lack of order and organization, it followed that in the distant future the world would irreversibly dissolve into an eternal state with no structure and no life. The world would still exist, but it would be a dead world. The pessimistic prediction of a global *heat death* gave rise to a 40-year-long cosmological debate that involved some of the leading philosophers, theologians, and social critics of the late Victorian era. Reluctant to deal with concepts as metaphysical as the universe and its entropy, almost no professional astronomers took part in what has been called the *entropic controversy*.

Opponents of the heat-death scenario not only found it absurd and philosophically offensive, but they also realized that it implied a universe with a beginning in time and might therefore suggest a supernatural creation. The controversy over the cosmological implications of the second law was an integral part of the more general ideological struggle associated with the rise of materialism, atheism, and socialism as alternatives to traditional ways of conceiving nature, culture, and society. Among the participants were the philosophers Herbert Spencer and Friedrich Nietzsche, and communist pioneers such as Friedrich Engels and Louis-Auguste Blanqui. Although the controversy was more ideological than scientific, it concerned the development of the universe as a whole and as such, it was truly cosmological. Scenarios somewhat similar to that of heat death are still parts of modern cosmological theory.

20th-Century Observations

Observational cosmology in the early part of the 20th century was much concerned with the enigmatic spiral nebulae and their relations to the

Milky Way system. Harlow Shapley (1885–1972) and some other leading astronomers held that the material universe was essentially made up of one enormous galaxy, our Milky Way, with the nebulae being located in the remote parts of the Milky Way system. Other astronomers were in favor of the idea of an *island universe* first suggested by the philosopher Immanuel Kant in a work of 1755. According to this hypothesis, the spiral nebulae were independent galaxies far away from the Milky Way and approximately of the same size and structure. The debate between the two rival views of the universe was only settled in 1925 when Edwin Hubble at the Mount Wilson Observatory succeeded in measuring the distance to the Andromeda Nebula to be about 1 million light-years. With Hubble's discovery, the island universe theory was vindicated, and the nebulae became recognized as galaxies.

A few years later, Hubble (1889–1953) turned to a spectroscopic research program to find a relationship between the distances of the galaxies (r) and their red shifts. In a famous paper of 1929, he established from observations a linear relationship between the two quantities. Interpreting the red shift as an apparent radial velocity (v), he found that $v = Hr$, in which the constant of proportionality H is known as the *Hubble constant*. For the most distant of the galaxies in his data set, Hubble found that it receded from the Earth with a velocity of about 1,000 km/s. Thus, it seemed that all spiral nebulae or galaxies moved away from the Earth with recessional velocities increasing with their distances. Reluctant to go beyond observations, however, Hubble did not conclude from the $v = Hr$ law that the universe expands or that the galaxies actually move apart from each other. In fact, he never unambiguously supported the idea of an expanding universe of which he is often considered the father. Although his paper of 1929 is today hailed as a revolution in cosmological thought, this was not how the cautious Hubble looked upon it. Until his death in 1953, at a time when the expanding universe had long been accepted, he remained a skeptic, arguing that a static universe was still a possibility that could not be ruled out observationally.

Static and Expanding Models of the Universe

Shortly after having introduced his general theory of relativity, a new theory of gravitation that would come to replace Newton's, Einstein applied it to the universe at large. His cosmological theory of 1917 assumed a finite yet unbounded space uniformly populated with stars. Although this was not possible within the framework of Newtonian physics and Euclidean geometry, the space of Einstein's universe was positively curved, like the surface of a sphere. The idea of non-Euclidean curved space goes back to the 1830s but was for a long time considered a mathematical curiosity of no relevance to astronomy or physics. Einstein was not the first to apply curved space to cosmology, but he was one of the first to do so and by far the most important.

In order to secure a static, noncollapsing universe, Einstein added to his gravitational field equations an antigravity term called the *cosmological constant*, which he found to be necessary but eventually abandoned. Being static, the closed Einstein universe had no beginning in time and no end. Throughout the 1920s, a small group of physicists, astronomers, and mathematicians investigated the Einstein model, which they found to be satisfactory but far from perfect. One of the problems was that it could not account for the mysterious nebular red shifts found by Slipher.

Einstein thought, erroneously, that his 1917 model was the only realistic solution to the cosmological field equations. Although the Russian physicist Alexander Friedman proved in 1922 that there were solutions corresponding to expanding and other evolutionary models, his brilliant insight was ignored. The same was the case with an important paper of 1927 written by the Belgian physicist and Catholic priest Georges Lemaître, who independently of Friedman predicted an expanding version of the Einstein universe. Moreover, Lemaître demonstrated that the red shifts of the spiral galaxies could be explained as an effect of the expansion of cosmic space. In fact, he derived theoretically the linear Hubble law 2 years before Hubble presented it as an empirical law. Alas, the works of Friedman and Lemaître

were only recognized in 1930, when Arthur Eddington in England, Willem de Sitter in The Netherlands, and a few other prominent astronomers concluded that the static universe had to be replaced by an expanding one. Only then was Hubble's red shift-distance relation from the previous year interpreted as evidence for an expanding universe.

By the mid-1930s, the expanding universe was accepted by a majority of professional astronomers and physicists, but outside the circles of avant-garde scientists, the theory met with considerable opposition. The notion of an expanding universe might seem contradictory, for how can the universe—which comprises everything—become bigger and bigger? If it expands, what does it expand into? A minority of astronomers wanted to retain the more comprehensible static universe and consequently developed alternative hypotheses to account for the Hubble relation on which the expanding universe theory relied. For example, the red shifts could be explained on the *tired light* assumption that galactic photons gradually lost energy on their journey through space. For a while, the static alternatives were taken seriously, but since about 1950 they have been generally considered outside of mainstream science.

Big Bang Versus Steady State

One might believe that with the acceptance of the expanding universe followed the acceptance of a universe with a beginning time, on the grounds that if the universe expands continually there must have been a time in the past at which its size was zero or close to zero. As illustrated by the historical development of cosmological theory, however, a universe of finite age is not necessarily a consequence of an expanding universe. For example, Lemaître's model of 1927 had no definite beginning but evolved slowly from a preexisting and potentially unstable spherical Einstein universe. The audacious hypothesis of a universe with an origin in a small and very compact state was introduced by Lemaître in a paper of 1931, which marks the beginning of *big bang* cosmology. The Belgian cosmologist suggested that many billion years ago, all matter in the universe was concentrated in a primordial, highly radioactive body with a density of the same

order as an atomic nucleus. Space and time did not yet exist, but only emerged with the radioactive explosion of the hypothetical proto-universe or *primeval atom*, as Lemaître called it.

This first big bang model attracted much public interest but was received with reservation by most astronomers and physicists, who either ignored it or rejected it as hopelessly speculative. The idea of a universe that had somehow been created a finite time ago was considered quite radical, and through the 1930s it was accepted by less than a handful of cosmologists. The reason was not only the conceptual novelty of Lemaître's primeval-atom hypothesis but also that there was no empirical evidence that clearly supported it. It was only in the late 1940s, with the pioneering work of the Russian-American nuclear physicist George Gamow, that the big bang hypothesis was developed into a viable physical theory of the early universe.

With little or no inspiration from Lemaître, Gamow assumed the initial state of the universe to consist of a hot and highly compressed soup of neutrons. From this scenario, he and his collaborators tried to calculate how nuclear reactions would produce hydrogen and helium nuclei, and eventually all the nuclei of the known chemical elements. Gamow hoped to explain the observed cosmic abundance of elements, but although he succeeded with respect to helium it proved impossible to account for the formation of heavier elements. Nonetheless, Gamow's theory of the early universe attracted much interest, if more from nuclear physicists than from astronomers. During the period from about 1948 to the early 1960s, Gamow was the chief exponent of what came to be known as the hot big bang model.

The basic idea of a universe evolving in accordance with the equations of general relativity and with a past completely different from the present state was not generally accepted in the 1950s. According to the rival theory proposed in 1948 by three young Britons (Fred Hoyle, Hermann Bondi, and Thomas Gold), the universe had existed in approximately the same state for an infinity of time and would remain in the same state for another infinity of time. On a large scale, the universe satisfied the *perfect cosmological principle* of uniformity with respect to both space and time. Since the three physicists accepted the expansion of space, they had to postulate continuous

creation of matter in order to keep the mass density of the universe constant.

The steady-state theory, as it was called, was fundamentally different from the class of evolution theories based on general relativity, whether of the big bang type or not. From a methodological point of view, it was attractive because it resulted in a number of sharp and testable predictions, but it was also controversial because of its postulate of matter creation which contradicted the authoritative law of energy conservation. The protracted controversy between the two rival views of the cosmos was primarily of a scientific nature, but it also involved a heavy dose of philosophical issues. Which of the two main conceptions of the universe was the most scientific, and according to what criteria? The Austrian-British philosopher Karl Popper's ideas about falsifiability were used by Bondi and others to suggest that the steady-state theory was methodologically superior to the class of big bang theories. Among the philosophers of science who entered the debate were prominent figures such as Stephen Toulmin, Rom Harré, Adolf Grünbaum, and Norwood Russell Hanson.

Religion too was an element in the controversy. Whereas opponents of the big bang theory intimated that it was apologetic, a scientific version of Genesis, in Christian circles steady-state cosmology was widely associated with atheism. If there were no beginning of the universe, it might seem that there was no need for a divine Creator. The rival cosmological views also became part of the ideological battlefield in the Cold War era. In the Soviet Union, and later also in communist China, big bang models came to be seen as politically suspect and contrary to the atheistic worldview of Marxism-Leninism. Nor was the steady-state theory considered acceptable on such grounds: Although it posited no cosmic beginning, the continuous creation of matter was rejected as politically incorrect by zealous Communist Party ideologues.

Relativistic Standard Cosmology

The steady-state theory led to several sharp predictions which could be tested observationally, such as the curvature of space and the rate of expansion of the universe. For a long time,

however, observations were too uncertain to rule out the theory defended by Hoyle and his allies. The new science of radio astronomy promised to fare better and generally suggested that we do not live in a steady-state universe. But again, the results from radio astronomy did not amount to an unambiguous refutation of the theory. The point of no return came only in the mid-1960s with the discovery of the cosmic microwave background radiation, which agreed perfectly with the big bang theory but was inconsistent with the rival steady-state theory.

As early as 1948, two of Gamow's collaborators, Ralph Alpher and Robert Herman, had predicted a cosmic background of photons created in the early phase of the universe, but it took until 1964 before the radiation was serendipitously discovered. Shortly thereafter, physicists at Princeton University realized that the microwave background provided decisive support for the big bang, whereas it was incompatible with the steady-state theory. At about the same time, measurements of the cosmic abundance of helium led to the same conclusion. Thus, by the late 1960s, the theory proposed by Hoyle, Bondi, and Gold was no longer a viable alternative to the now victorious big bang theory based on general relativity and nuclear physics.

A few years after the discovery of the cosmic background radiation, there emerged a consensus view that the relativistic big bang theory was the one and only foundation of scientific cosmology. Although alternatives did not disappear, they were marginalized. At the same time that cosmology became cognitively institutionalized, it achieved a social institutionalization that made the subject a full-time occupation rather than a part-time hobby for astronomers and physicists. The field was increasingly integrated into university departments, and courses and textbooks on cosmology became common. For the first time, students were taught standard cosmology and brought up in a research tradition with a shared heritage. The publication in 1970 of James Peebles's *Physical Cosmology*, the first-ever textbook with this title, marked the beginning of a new and progressive development.

Whereas cosmology had earlier been a subfield sandwiched between astronomy and physics, it now aspired to become a scientific discipline in its

own right. With the professionalization and status as a respected field of science, there followed the exclusion of philosophers and amateur scientists, a development in stark contrast to the situation during the 1950s. Moreover, the national differences that had characterized earlier cosmology—big bang theory being American and steady-state theory British—also disappeared. By 1970, cosmology had become truly international, if not yet forming a scientific community.

Dark Energy and Other Surprises

The cosmological standard model in the 2020s is still a big bang universe governed by general relativity, as it was in the 1970s. And yet our present understanding of the universe differs significantly from the model favored 50 years ago, which was a universe with decreasing expansion rate, zero cosmological constant, and made up of ordinary matter in the form of atoms and their constituent particles. The cosmological constant (symbol Λ , lambda), which Einstein introduced in his static model of 1917, was since about 1930 considered to be a mistake (i.e., $\Lambda = 0$) and ignored by most cosmologists until it unexpectedly returned at the end of the century.

Observations from 1998 of distant supernovas showed that the rate of cosmic expansion is not decreasing, but increasing. We live in an accelerating universe. To account for the blowing up of space revealed by observations, physicists reintroduced the repulsive cosmological constant as a measure of the energy content of empty space, so-called *dark energy*. Although this strange form of self-generating energy has not been directly observed, it is generally believed that it fills all of space and dominates the fate of the universe. If the picture is correct, dark energy will cause the universe to expand indefinitely and at a still greater rate, meaning that in the far future it will consist almost entirely of lifeless empty space filled with dark energy. The heat-death scenario discussed in the late 19th century has returned, if in a different form.

As the modern universe is characterized by a large amount of dark energy, so is it characterized by an almost equally large amount of so-called *dark matter*, a concept going back to the 1930s but substantiated only some 40 years later from

observations of the rotation of galaxies and other evidence. As the name indicates, dark matter is a hypothetical form of invisible matter which acts gravitationally but not electromagnetically. Whatever dark matter is, it is entirely different from the ordinary matter consisting of protons, neutrons, and electrons, which makes up only 5% of the mass of the universe. The most recent data from 2019 indicate a mass distribution of 71% dark energy and 24% dark matter. Yet another discovery of momentous cosmological importance concerns the cosmic microwave background, which since 1989 has been explored by means of satellite experiments. The data from these and other experiments have revealed small variations in the energy density of the otherwise uniform radiation, from which cosmologists have derived a wealth of information about the very early universe. The many measurements from satellites and other experiments have led to a surprisingly precise determination of the age of the universe, which by 2019 was believed to be 13.787 billion years, with an uncertainty of only 0.020 billion years.

The Consensus Model and Its Critics

The standard model established in the early years of the new millennium is also known as the Λ CDM *model* because it relies on the cosmological constant (Λ) and cold dark matter (CDM), the term *cold* meaning that the hypothetical dark matter particles move relatively slowly. The model has been very successful insofar as it has an impressive explanatory and predictive record, and for this reason it is the accepted framework for mainstream work in theoretical and observational cosmology. It is generally admitted, however, that there are flaws or at least unsolved problems in the Λ CDM theory. One of them concerns dark energy, namely that the observed value of the cosmological constant differs dramatically from the one calculated on the basis of quantum mechanics. The persistent failure in detecting the particles constituting dark matter is another serious problem.

Whereas the standard model assumes the universe to have come into existence with the big bang, some physicists argue from theories of quantum gravity, such as string theory, that our universe was preceded by a *pre-big bang universe*

stretching infinitely back in time. The collapse of this earlier universe gave rise to a *big crunch* out of which the big bang emerged. These theories are to some extent mathematical speculations, but in principle and possibly also in practice they are empirically testable. The same is the case with the controversial and much-discussed *multiverse* hypothesis from about 1990, the claim that there exist a huge number of other universes in addition to the one we inhabit. The other universes may be causally separated from ours and may have different natural constants and physical laws, which makes the multiverse hypothesis even more controversial.

Although the standard Λ CDM model functions as a kind of paradigm in the sense meant by the historian of science Thomas Kuhn, it is not beyond criticism. While it has received little attention from philosophers and sociologists of science, a minority of physicists and astronomers have expressed strong dissatisfaction with the current state of cosmology. One of them is the prominent South African cosmologist George Ellis, who has warned against metaphysical tendencies in parts of modern cosmology such as the belief in a spatially infinite universe. According to Ellis, infinities do not belong to science. Other critics, among them the astrophysicists Robert Sanders and David Merritt, focus on what they consider to be the unfounded postulate of dark matter, a central element in the standard model. In part from philosophical positions, they suggest that the current cosmological paradigm is a dogma that prevents innovative thinking and a necessary revision of how to understand the universe. A paradigm shift is therefore needed. However, the critical voices have not seriously shattered the cosmological community's general confidence in the standard Λ CDM model.

Helge Kragh

See also Astronomy; Falsifiability; Geometry, Non-Euclidean; Physics, 20th Century; Physics, Gravitation Theory; Worldview

Further Readings

Bertotti, B., Balbinot, R., Bergia, S., & Messina, A. (Eds.). (1990). *Modern cosmology in retrospect*. Cambridge University Press.

- Coles, P. (Ed.) (2001). *The Routledge companion to the new cosmology*. Routledge.
- Crowe, M. J. (1994). *Modern theories of the universe: From Herschel to Hubble*. Dover.
- Gribbin, J. (1997). *Companion to the cosmos*. Phoenix Giant.
- Harwit, M. (2013). *In search of the true universe: The tools, shaping, and cost of cosmological thought*. Cambridge University Press.
- Hetherington, N. S. (Ed.) (1993). *Encyclopedia of cosmology: Historical, philosophical, and scientific foundations of modern cosmology*. Garland.
- Kragh, H. (1996). *Cosmology and controversy: The historical development of two theories of the universe*. Princeton University Press.
- Kragh, H. (2007). *Conceptions of cosmos. From myths to the accelerating universe: A history of cosmology*. Oxford University Press.
- Kragh, H., & Halvorsen, H. (2013). *Theism and physical cosmology*. In C. Taliaferro (Ed.), *The Routledge companion to theism* (pp. 241–255). Routledge.
- Kragh, H., & Longair, M. S. (Eds.) (2019). *The Oxford handbook of the history of modern cosmology*. Oxford University Press.
- Munitz, M. K. (1986). *Cosmic understanding: Philosophy and science of the universe*. Princeton University Press.
- North, J. (1990). *The measure of the universe: A history of modern cosmology*. Dover.
- Peebles, P. J. E. (2020). *Cosmology's century: An inside history of our modern understanding of the universe*. Princeton University Press.
- Sanders, R. H. (2016). *Deconstructing cosmology*. Cambridge University Press.
- Smith, R. W. (1982). *The expanding universe: Astronomy's "great debate" 1900–1931*. Cambridge University Press.

CYBERNETICS, 20TH CENTURY

At the beginning, cybernetics was a wide-ranging interdisciplinary effort spurred on by the mathematician and philosopher Norbert Wiener and a group of scientific colleagues with the intent to study communication, feedback, and control in biological, physical, engineering, and social systems. Feedback and circular causality associated with it was key to the investigations of this group so that the concept moved along from the simple

aspects of engineering regularities such as thermostats and regulators for steam engines and other mechanisms to the subtleties of the feedbacks in muscle/eye coordination in organisms, and the deeper subtleties of the regulation of speech and logic and cognition in the act of its production. Starting with the development of core themes in the emerging field of cybernetics, the entry continues with an account of the development of mathematical models and a discussion of reference and self-reference and their connection with language. An explanation of Gödel's Theorem follows, and the entry concludes with discussion of autopoiesis and the self-producing nature of living systems.

Key Themes

The theme of feedback and circularity widens and deepens as it is studied by cyberneticians and encompasses systems maintaining their structure in terms of their own action in a world of change. The original Macy Conferences on Circular Causality included all these themes, and with the inclusion of mathematicians of the caliber of Wiener and John von Neumann, the discussions ranged over the implications of computers, Turing machines, self-reproducing automata, information, logic, the beginnings of automation, and the social consequences of automation. Warren McCulloch and Walter Pitts joined the group and lived in the long wonder of how it can be that mind, logic, and thought can be interwound with the workings of the brain and with the properties of networks and circularly interconnected systems. McCulloch wrote papers with titles like "What is a Man that he may know a Number and what is a Number that it can be known by a Man?" With the inclusion of the anthropologists Gregory Bateson and Margaret Mead, the discussion extended itself to questions of how individuals and societies worked together and how the world/worldview of an individual is shaped by their society and how their society is shaped by the individual. This theme of circularity becomes the always-changing basis for the depths of cybernetic study.

The original discussions of cybernetics were deeply related to the meaning of that term as embodied in the Norbert Wiener's title for his book *Cybernetics—or Control and Communication in*

the Animal and in the Machine. The book was published in 1948 and was Wiener's summary of interdisciplinary discussions that had begun 10 years earlier, prior to World War II. These discussions were influenced by the war and by such problems as understanding computers as they were emerging, analog and digital, and understanding how to predict and control aircraft and navigation processes. The discussions moved across fields including physics, biology, and engineering. The initial direct problems of steering a gun, a ship, or one's own body directly or in the operation of a prosthesis led Wiener and his group to study the properties of feedback and the circular processes associated with it and to hypothesize this area of feedback and circular causality as fundamental in biology, engineering, and human planning. Concomitant with feedback in this sense is the sort of communication that supports it, and so large themes of communication, information and stability, instabilities and the possibility of homeostasis arrive in the core of cybernetics. The interdisciplinary discussion was fed by streams from all its participants and did not stay at the level of mechanical or direct physical-biological actions.

Pioneers in Cybernetic Thinking

The question of the influence of communication and feedback for the action of societies and groups of individuals was just as important and indeed was and is a central concern for cybernetics. Once communication and information became a theme, questions of quantitative measurement arose, and one finds the work of Claude Shannon speaking to that. Questions of mind arise, and one finds the work of McCulloch and Pitts speaking to that. As soon as the social level enters the discussion, the role of the observer is present and can be found in Wiener's discussions of language and society in his book. It was a major theme for Wiener to see the possibilities of automation and its influence on society. Again the influence of the war years is crucial in the history of these developments. Consider the role of control and communication in relation to secret communications, the coding in such devices as the Enigma Machine and the key role of early computers, recursion, and the analysis of languages in the decoding, all for the

sake of prediction and control in the war effort. At the same time, the message structures of nationalistic propaganda were in battle with these same forces, and nuclear technology was under development—all for the sake of manifest levels of control and survival. There is no way to separate technology from the social structures that it affects. The themes of feedback, communication, and control are integral to the understanding of interactions at all levels. All these themes are now upon us again and always as we watch and participate in the influence of worldwide computer networks, the rise of so-called artificial intelligence and the continuing need to find ways through the complexities of human relationships and decisions in the worlds created by our own recursions.

In the wake of the original Macy Conferences, Heinz von Foerster became an eloquent spokesperson for the themes of cybernetics. Others, such as Ross Ashby, arose and wrote books on the principles of the subject. New explorers such as Stafford Beer, Gordon Pask, Humberto Maturana, Francisco Varela, Ernst von Glasersfeld, Ranulph Glanville, and Annetta Pedretti came forward. The field continues its wide thematic approach. Much of modern thought, computer science, systems thinking, complexity thinking, studies in chaos theory, logic, and ideas about reflexivity and interdependence have their origins in cybernetics.

Second-Order Cybernetics

A significant shift in the definition of cybernetics occurred with Mead's injunction that the methods of cybernetics should be applied to cybernetics itself, and with von Foerster's naming of this new cybernetics as *second-order cybernetics*. Second order was to be the cybernetics of observing systems, where the system was its own observer, its own cognizer, just as a human being can be an observer of themselves, and just as language can discuss and refer to language itself.

At the concept of the second order, we arrive at closure. It should be noted that the cybernetics of the second order had been practiced by cyberneticians from the beginning. Nevertheless, it is a significant shift of viewpoint to name this way of thinking and being. Having come to second-order

cybernetics and its inevitable reflexivity, we are now prepared to explore the many ramifications of the cybernetic concept.

One way to understand the emergence of second-order cybernetics is to see it as completing a larger circularity, a feedback loop that includes Meaning and Form (form, syntax, logic, scientific explanation) in one whole:

Syntax gives rise to meaning.

Meaning gives rise to syntax.

In this characterization, the observer holds meaning and relationship, and the participation of individuals with one another and with their activities. These activities have form, pattern, syntax, and transferability that allow explanation and the possibility of interdisciplinary comparison. Meaning can always ask more and demand a larger context. There is a sense of the whole that consists in meaning and form in dialogue. It is at the level of such wholes that one can ask about the meaning of meaning or the nature of number or the role of consciousness in systems. This thematic relationship has always been present in cybernetics, but the emergence of a conscious statement of second-order cybernetics as a form of cybernetics has made the difference between cybernetics as a commentary on science, engineering and computation and social structure, and cybernetics as a true transdisciplinary activity.

It is at the level of shared pattern (the level of concept) that we find our commonality and the possibility of mutual participation that is key for cybernetics and other transdisciplinary studies. Just so it is at the level of shared and understood language that we find that commonality. To begin to understand language is to begin to understand the nature of pattern and the role of cybernetics.

Mathematics, Language, and Cybernetics

It must be pointed out that mathematics is another kind of languaging, letting an observer play with symbols on paper without themselves seeming to appear on that paper. In the course of that play, the observer is involved and not separated from the actions in the seemingly separate world of the paper. The meaning that this activity generates

weaves a deeper world where the observer and their actions with paper or other media are inseparable from themselves.

The sorts of mathematical models that have been and are being used in cybernetics are as follows: mathematical logic, lambda calculus, Laws of Form, reflexive domains, Turing machines, mathematics of recursion, logical networks (as in the work of McCulloch and Pitts), general circularly interconnected networks that can model systems of all kinds—including graphs and general systems of graphs, and category theory. Recursions and cellular automata became widely studied after the 1960s, and some of these themes are of particular importance in cybernetics.

For example, Mitchell Feigenbaum discovered that there were universal properties in the real line recursion $x' = kx(1 - x)$. That such simple examples of feedback can have a wide applicability and increase our understanding of mathematics and nature is a puzzle for those interested in cybernetics from its ground up.

The properties of recursions such as $x' = kx(1 - x)$ are noteworthy *because this phenomenon places us squarely in between interpreting cybernetics in the second order and interpreting it in the first order*. If one looks into the mathematical and computational properties of a recursion, it may appear that one is only examining a technicality using a standard Western objective epistemology. This is not so, in that there is no necessary link between examining the mathematics of a recursion and one's assumptions about the nature of mind and the mind of nature. Nevertheless, one has the situation that the geometry of processes like $kx(1 - x)$ reveal great depths unsuspected prior to the mathematical experiment. Mathematics advances by observing its own productions. One way to understand this situation is to examine how, in handling recursions, the observer takes a role. Then we can look at intensities like $kx(1 - x)$ and ask about the role of the observer in such a study. In so doing, we confront the roles of awareness, distinction, and the participation of the observer in the production of reality.

We begin in language. We begin in the situation where language and metalanguage, language about language, are one.

Language = MetaLanguage

To live in thought and language is to live in dialectic and paradox, circularity, and feedback. This is the cybernetic beginning. From here it is natural to examine all that might lead to clarity and logic and science. How did it happen that we managed the regularities that we have actually managed? What is the role of reason in this fundamental circularity? From this reflexive place, we can see the source of cybernetic questions from beginning to end. And we can see how mathematical and symbolic structures play a key role in providing the glue that binds discursive, logical, and quantitative thoughts.

Cybernetics has a long history from interdisciplinary discussions of many instances of feedback, communication, and control to becoming self-conscious of its own aims and being willing to apply itself to itself. We have looked at the role of language, sign, and symbol for the understanding of understanding, the describing of describing. These produce domains of action that involve the creation of new language, new mathematics, and new approaches to praxis.

Distinction, Awareness, and Self-Reference

It is possible for a sign to stand for itself.

For example, a mark $< >$ can be regarded as a sign for the distinction that it makes between its inside and its outside.

The mark is a sign that stands for itself.

In his seminal work *Laws of Form*, G. Spencer-Brown takes the self-referential mark of distinction as a starting point and the equations

$$<><> = <>$$

$$<<>> =$$

as the expression of a calculus of indications. The first equation can be interpreted as the redundancy of calling the name of the mark by itself. If I wear a name tag, it is not necessary. The second equation involves regarding the sign $<>$ as an act of distinction or crossing. To cross from the unmarked state achieves the marked state: $<> = <>$. To cross from the marked state achieves the unmarked state $<<>> =$. The calculus of indications of Spencer-Brown is a

minimal formalism in the enactment of meaning and syntax.

Contraction of reference leads to the expansion of awareness.

Language, Reference, and Self-Reference

We can discuss describing through a simple language involving the digits 1, 2, 3 and their meanings as one, two, three. Consider the pattern of numbers below:

1

11

21

1211

111221

Each line is a description of the previous line. To see this, read the lines aloud. The second line says "one one" and that is a description of the first line. The third line says "two ones" and that is a description of the second line. The next line says "one two, one one," then "one one, one two, two ones," and so on. The full alphabet for this recursion is the set of numerals {1, 2, 3}, and these are alternately signs and elements of the description of a pattern. This *audioactive sequence* was extensively investigated by John Horton Conway and has many mathematical properties.

In the audioactive formalism, we have that $2x$ describes xx . Thus, 22 describes 22, and we arrive at self-reference in this small language. We can write $a \rightarrow b$ where the arrow denotes that " a describes b ." Then we have $2x \rightarrow xx$ and letting $x = 2$, we have $22 \rightarrow 22$.

We could begin with $gx \rightarrow xx$ and obtain $gg \rightarrow gg$ by substitution, but there is no given meaning for g or for the arrow $gx \rightarrow xx$. By generalizing the meaningful pattern $2x \rightarrow xx$, to an apparently meaningless pattern $gx \rightarrow xx$ (from meaning to syntax) we enter a wider domain of possible meanings (hence from syntax to wider meaning).

This is particularly so with the work of Alonzo Church and Haskell Curry on the Lambda calculus, since they generalize even further to the pattern $gx = F(xx)$ and then obtain a fixed point for

F by the substitution $gg = F(gg)$. In fact, they were motivated in this endeavor by the special case of $Rx = \sim xx$ where $\sim$ denotes negation in logic and AB denotes that “B is a member of A.” Then $Rx = \sim xx$ is interpreted to mean that *x is a member of R exactly when x is not a member of x*. This is the Russell set. We obtain $RR = \sim RR$, meaning that R is a member of R exactly when R is not a member of R. This is the Russell Paradox. The paradox asserts that negation can have a fixed point. In a cybernetic context, this is a viable possibility.

The route to self-reference we are discussing widens to a process I call the *indicative shift*.

Let

$$b \rightarrow B$$

denote a *reference of b to B*. You can take b to be a name for B.

The *shift* of this reference is denoted by

$$\#b \rightarrow Bb.$$

I am introduced to Mister B and my host says to me, please meet “b.” Being an attentive guest, I make an association of the name b with the appearance B and put them together in my mind. The next time I meet B, he appears to me as a Bb in the sense that the name comes right along with him. In my imagination, B has the name tag b on his lapel. I also have that name stored with a marker #b (made explicitly here but usually unsaid). The name #b is now pointing to my amalgam of B with his name. To examine your own process in this regard, think of the times you have encountered someone whose name you have forgotten. You find yourself attempting to reconstruct the links we have just described.

What does the indicative shift have to do with self-reference? Let M be the name of the shift. Then

$$M \rightarrow \#.$$

Shifting, we find that

$$\#M \rightarrow \#M.$$

The meta-name (#M) of the name of the shift refers to itself.

We rewrite:

I am the meta-name of my meta-naming process. and find a relative of the Heinz von Foerster sentence

“I am the observed relation between myself and observing myself.”

Begin the indicative shift at the most elemental point.

$$\rightarrow$$

This is an arrow from nothing to nothing. Nothing stands for nothing. The arrow is prior to names.

Apply the shift.

$$\# \rightarrow$$

There was nothing to shift.

Apply the shift again.

$$\#\# \rightarrow \#$$

And again.

$$\#\#\# \rightarrow \#\#\#.$$

Thus, we have

$$\rightarrow$$

$$\# \rightarrow$$

$$\#\# \rightarrow \#$$

$$\#\#\# \rightarrow \#\#\#.$$

At the third departure from the void, we find that self-reference has occurred.

Meaning arises from syntax in our understanding of the process that it connotes.

In the beginning there was no name.

The shift became the name of nothing.

The shift of the name of nothing became the name of the shift.

The shift of the name of the shift is its own name.

Here begins a cybernetics of the acts of distinction.

The self-reference of the indicative shift is more subtle than Two Twos. It involves the action of observing and it shows how the act of observing (and naming) turns around and names itself. We are aware of the realm of meaning for the observer and here we have begun steps into a syntax for the observer. The second order cybernetic loop of Meaning and Syntax must continue

into great richness before we can be willing to say that we have a cybernetics of observing systems.

May that happen in the course of time.

The Gödelian Shift

The indicative shift occurs in reference and naming in language and communication. A shift of this shift takes us to the key structures of Gödel's incompleteness theorem.

Kurt Gödel proved that no consistent formal system rich enough to handle number theory could be complete. Gödel showed that there are true theorems about numbers that the given formal system cannot prove.

Gödel produced a sentence that encoded a denial of its own provability. He devised a method to code each formula F in his system with a number $g = g(F)$ (the Gödel number) so that the formula could be uniquely decoded from its number. I will write $g \rightarrow F$ to denote that " g is the Gödel number of F ."

Now suppose that $F(u)$ is a formula with a free variable u .

For example $F(u)$ could be " u is a prime number." Let g be the Gödel number of $F(u)$. Then we can substitute g into $F(u)$ to obtain $F(g)$. This is a new formula with a new Gödel number, call it $\#g$. Then we have $\#g \rightarrow F(g)$. This is the "Gödelian indicative shift" of $g \rightarrow F(u)$.

$$g \rightarrow F(u).$$

$$\#g \rightarrow F(g).$$

Now the function assigning the number $\#g$ to the number g is an algorithm about numbers, just the sort of thing that Gödel's formal system L can talk about. Thus, we can have $\#$ as an element in the language of L .

Let $B(u)$ be a statement in L that asserts the provability of the statement with Gödel number u . Then $\sim B(u)$ asserts the unprovability of the statement with Gödel number u . And furthermore, we have a Gödel number for $\sim B(\#u)$, the statement that the formula with Gödel number $\#u$ is not a provable formula:

$$g \rightarrow \sim B(\#u).$$

Then making the shift, we have

$$\#g \rightarrow \sim B(\#g).$$

This shows that $\sim B(\#g)$ asserts the unprovability of the formula with Gödel number $\#g$. But that formula is $\sim B(\#g)$! This means that $\sim B(\#g)$ asserts its own unprovability.

If L could prove this formula, then L would be inconsistent.

We assume that L is consistent and conclude that it cannot make the proof.

But that is exactly what the formula says, and so $\sim B(\#g)$ is true but not provable in L .

We have sketched the proof of Gödel's incompleteness theorem and as you see, the Gödelian indicative shift is the key mechanism whereby self-reference is achieved to obtain a theorem that asserts its own unprovability. The self-reference is accomplished via the coding from texts to Gödel numbers, and so is protected from paradox. Here the pendulum swings from the meaningful arena of the indicative shift in the naming processes of ordinary language to the highly syntactical regions of formal systems.

The Gödel Theorem deserves to be seen as second-order cybernetics. Its content requires an observer and their understanding. The Gödel Theorem requires the rational comprehension of an observer of the formalism. In working with the final shifted equation

$$\#g \rightarrow \sim B(\#g)$$

we stand outside the formal system, understanding the meaning of the reference that makes $\sim B(\#g)$ state its own unprovability. We prove that L will produce a contradiction within its own syntax if L should produce a demonstration of $\sim B(\#g)$. We reason structurally about L through our relationship with L . We have access to the properties of the shift and an ability to reason about it that is not available to L .

We begin to understand how we as observers can be in intimate relation with a formal system. We can go forward into the classical and deep questions of the relationship of ourselves and machines (formalities, syntax). We return as always to the understanding that cybernetics studies the feedback loop of meaning and syntax and see this relationship anew.

Autopoiesis and the Meaning of Living Form

It has long been understood in cybernetics that one should think of living systems and indeed larger organizations such as societies, worlds, galaxies, and the universe herself as producing themselves out of their own productions in the contexts of those environments appropriate to such actions.

Even our simplest example of Two Twos produces itself out its own production, where that production is the meaning associated with the syntax. But in autopoietic models, we can begin to see how self-sustaining systems can arise from the circular interactions of a (e.g., molecular biological or social) given substrate. Circularity is a natural condition of process.

To be produced out of one's own productions is to be an *eigenform* of oneself. Look at how different thinkers have handled this notion and what wisdom they have given us. McCulloch wanted to know how the brain could find number and how number could find the brain. How does the brain produce concept from itself and from its environment? How does concept influence the brain? McCulloch and Pitts did pioneering work in taking symbolic logic to a domain for modeled neural processes. Their work comes down to the present day in neural nets and machine learning and all the matters of logic and computing. The fundamental questions are as mysterious as ever. Conceptual calculus is not identical to its formalism any more than second-order cybernetics is identical to the study of feedback and machines.

Others, such as Gordon Pask, have investigated language with geometry and networks and topological possibilities. Sensitive intelligences such as Humberto Maturana saw living processes directly in terms of the worlds those processes created in the course of producing themselves. An organism producing itself is continually reacting not to an outside world but to the perturbations of its context and stability by what we observers perceive to be its outside world. We as observers of that organism see more and see less than the organism itself. We see structure and limitation and properties of the environment. We have not the sight to be the world of that plant or animal. We have intuition from our own experience of being, but

that experience does not fully cross the gap between one of us and another. And so it is that cybernetics shows rather than directly says (compare Wittgenstein and Spencer-Brown) what can bridge the unbridgeable gaps of communication that are needed for us to produce ourselves from ourselves.

Louis H. Kauffman

See also Actor-Network Theory; Communication Theory, Technical Overview; Information Theory

Further Readings

- Ashby, W. R. (1956). *An introduction to cybernetics*. Chapman & Hall.
- Bateson, G. (1979). *Mind and nature, a necessary unity*. E. P. Dutton.
- Beer, S. (1972). *Brain of the firm*. Penguin Press.
- Conway, J. H. (1987). The weird and wonderful chemistry of radioactive decay. In T. M. Cover & B. Gopinath (Eds.), *Open problems in communication and computation* (pp. 173–188). Springer-Verlag.
- Glanville, R. (2009). A (cybernetic) musing: Black boxes. *Cybernetics and Human Knowing*, 16(1/2), 153–167.
- Feigenbaum, M. J. (1976). Universality in complex discrete dynamics, Los Alamos Theoretical Division Annual Report 1975–1976.
- Isaacson, J. (1981). Autonomic String Manipulation System, U.S. Patent Number 4286330, August 25, 1981.
- Kauffman, L. H. (1987). Self-reference and recursive forms. *Journal of Social and Biological Structures*, 10, 53–72.
- Kauffman, L. H. (2005). Eigenform. *Kybernetes*, 34, 129–150.
- Kauffman, L. H. (2012). Categorical pairs and the indicative shift. *Applied Mathematics and Computation*, 218, 7989–8004.
- Los Alamos Scientific Laboratory. (1975–1976). *Theoretical division annual report*. Scientific Laboratory.
- Maturana, H., & Varela, F. ([1973] 1980). Autopoiesis and cognition: The realization of the living. In R. S. Cohen & M. W. Wartofsky (Eds.), *Boston studies in the Philosophy of Science* (Vol. 42). D. Reidel.
- McCulloch, W. S. (1989). *Embodiments of mind*. MIT Press.
- Mead, M. (1968). The cybernetics of cybernetics. In H. von Foerster, J. D. White, L. J. Peterson, & J. K.

- Russell (Eds.), *Purposive systems* (pp. 1–11). Spartan Books.
- Pask, G. (1961). *An approach to cybernetics*. Hutchison.
- Pedretti, A. (1981). The cybernetics of language. PhD Thesis, Brunel University, Department of Cybernetics, Uxbridge.
- Spencer-Brown, G. (1969). *Laws of form*. George Allen and Unwin.
- Varela, F. G., Maturana, H. R., & Uribe, R. (1974). Autopoiesis: The organization of living systems, its characterization and a model. *Biosystems*, 5, 187–196.
- Varela, F. J. (1976). Not one not two. *The Coevolution Quarterly*, 12, 62–67. <https://cepa.info/2055>
- Varela, F. J. (1979). *Principles of biological autonomy*. North Holland.
- von Foerster, H. (2003). *Understanding understanding: Essays on cybernetics and cognition* (pp. 1–19). Springer. (Originally published as Observing Systems, The Systems Inquiry Series, Intersystems Publications [1981]).
- von Foerster, H., Mead, M., & Teuber, H. L. (Eds.). (1951). Cybernetics: Circular causal and feedback mechanisms in biological and social systems. Transactions of the seventh conference. Josiah Macy, Jr. Foundation.
- von Foerster, H., J. D. White, L. J. Peterson, & J. K. Russell (Eds.), *Purposive systems*. New York: Spartan Books.
- von Glasersfeld, E. (2001). The radical constructivist view of scienceThe radical constructivist view of science. In A. Riegler (Ed.), *Foundations of Science, special issue on “The Impact of Radical Constructivism on Science,”* 6(1–3), 31–43.
- Wiener, N. (1948). *Cybernetics. Or control and communication in the animal and the machine* (2nd rev. ed. 1961). MIT Press.
- Wittgenstein, L. (1922). *Tractatus logico-philosophicus*. Harcourt, Brace and Company.

D

DATA AND PHENOMENA

Scientists often want to check whether their theories hold up in the face of results gleaned from observations and experiments. A natural thought might be to assume that when scientists check their theories in this way, they check their theories against the empirical data produced by observations and experiments. If theoretical predictions match the empirical data, then—one might think—that is good news for the viability of the theory in question, whereas if theoretical predictions conflict with the empirical data, that counts against the truth or adequacy of the theory.

James Bogen and James Woodward argued in 1988, however, that this natural thought is mistaken. Scientists do not check theories against empirical data directly, nor should they. Data are idiosyncratic in the sense that they are produced by the contribution of a variety of causal factors, including minute and practically untraceable influences besides the primary subject of research. High-level scientific theories, Bogen and Woodward argued, are not invoked to predict or explain the minute variations in data, but rather the stable and repeatable characteristics of phenomena that can be separated from irrelevant idiosyncrasies by data-to-phenomena reasoning. This entry explores the issues implied by that argument for the relationship between data and phenomena.

Consider a simple experiment to determine whether a coin is fair. The experimenter conducts

one hundred trials by flipping the coin and recording *heads* or *tails* for each flip. These records are the data. However, what is useful for checking whether the coin is fair is a summary statistic, for instance, the mean of *heads* throughout the hundred trials. A similar insight applies to significantly more complicated data. Thus, Bogen and Woodward argued that the sort of techniques and methods of inferential reasoning that scientists use to transform data into claims about phenomena are of significant interest for the epistemology of science in practice.

Explaining the Phenomena

Bogen and Woodward introduced an influential distinction between data and phenomena. For them, data provide evidence for the existence of phenomena. In contrast, phenomena provide evidence for theories. In their canonical 1988 paper, Bogen and Woodward emphasized that data cannot typically be predicted or explained by theory, whereas phenomena can. For the coin flip experiment, the hypothesis that the coin is fair neither predicts nor explains the outcome of any individual trial, but it does predict and explain the average number of heads over a long sequence of trials. Bogen and Woodward mentioned more sophisticated and realistic scientific examples to illustrate their distinction between data and phenomena. Their examples of data included photographs and records of a variety of detectors, while their examples of phenomena included types of

nuclear reactions and certain aspects of human cognition.

Bogen and Woodward discussed a simple example involving the melting point of lead. Repeated, sufficiently precise thermometer readings of a sample of lead at its melting point would produce measurements that differ slightly from one another—the data would have some scatter. Given certain assumptions about the causes of the scatter, one can reasonably estimate the true temperature of lead's melting point. As in the coin flip experiment, the mean value of the temperature readings is what is ultimately of interest for determining the melting point of lead—a value that Bogen and Woodward stressed may not actually coincide exactly with any of the individual readings. Scientific theories, they argued, provide high-level, systematic explanations only of phenomena and not of data. To explain the value of a data point, or of the values of a scattered data set, would seem to require a quixotic accounting of highly particular causal minutia. Typically, such explanations cannot be provided in sufficient causal mechanistic detail to count as *systematic* in the sense Bogen and Woodward intended. Regardless, such explanations are not what scientists usually seek to explain or predict with their theories. In contrast, to the capriciousness of data, phenomena exhibit stability and generality.

Challenges

Since Bogen and Woodward's original 1988 publication, other philosophers of science have raised several challenges to their arguments. A handful of these contentions are discussed below.

Portability of Data, Locality of Claims About Phenomena

Sabina Leonelli has invoked database curation practices in biology in order to challenge two main claims that she attributes to Bogen and Woodward. The first claim that Leonelli challenges is the idea that data are local in a manner that claims about phenomena are not. Since data are produced by particular causal processes that are specific to the context in which they were produced, one might think that they cannot directly have general evidential value. Rather, following

Bogen and Woodward, data serve as evidence for claims about phenomena, which in turn serve as evidence for general scientific theories. Their idiosyncrasies make particular data different enough across contexts to render them local in contrast to the stability and repeatability of phenomena. The particularities of a single temperature measurement may never be exactly repeated, but the melting point of lead itself is stable. Leonelli challenges this claim by arguing that suitably prepared data can travel beyond the specific context in which they were originally produced and indeed be put to evidential use elsewhere. For instance, in the biological sciences, data collected on a particular model organism may be useful to many different researchers focusing on different phenomena that can be studied via that model organism. By carefully extracting, labeling, and making data available for other researchers to search, database curators can help data *travel* and find new applications in diverse epistemic contexts. Thus, Leonelli argues that data need not always remain solely within the context of their provenance. So, while the idiosyncrasies of a particular datum may depend on the context-sensitive details of its provenance, suitably assisted scientists evidently sometimes repurpose data in diverse contexts without calling upon claims about phenomena as intermediaries.

The second claim that Leonelli challenges concerns the purported generality of claims about phenomena. She argues that in biological sciences claims about phenomena can be highly context dependent, and like data, require careful packaging in order to be put to use more broadly. For example, Leonelli highlights the diversity of crucial concepts and terminology that biologists use to characterize their phenomena of interest among adjacent research contexts. She describes how key terms such as *bud*, *pathogen*, and *cell wall* have different meanings to different research communities, and how the same phenomenon or entity (such as a virus particle external to the cell) can be associated with different terms. Leonelli argues that database curators need to also take such context-sensitive variety on the side of phenomenon identification and characterization into account when facilitating database searches and use. Claims about phenomena, Leonelli argues, are therefore not always *nonlocal* in the manner

that might be suggested by Bogen and Woodward. Leonelli agrees with Bogen and Woodward that claims about phenomena play a special kind of evidential role for which data are unsuitable, but she attributes this to their formulation as propositions (as opposed to the *marks* of data, like photographs or chart recordings) rather than to their special nonlocality.

Explaining Data

Greg Lusk has argued that Bogen and Woodward's insistence that scientists do not, and ought not, provide systematic theoretical explanations for idiosyncratic data misses interesting cases from science in practice that, he argues, do just that. According to Lusk, scientists do sometimes explain data as a way of identifying phenomena. His primary example involves a method that atmospheric scientists use to infer the concentrations of various gas species in the atmosphere. Electromagnetic radiation passing through the layers of Earth's atmosphere is absorbed by these gases in characteristic ways. Scientists collect data from that radiation on Earth's surface and process the data in such a way as to reveal the patterns the missing radiation leaves in the overall spectrum of the radiation. The phenomenon that these scientists are ultimately interested in is the distribution of atmospheric gases. To make empirically informed inferences about the distribution of gases from the spectra, they face a so-called inverse problem. Some gas species can interfere with characteristic absorption patterns of another species of gas. Without such interference, one might hope to infer the presence of a particular gas species in the atmosphere by noting the presence of its unique absorption pattern in the spectrum. However, the interference from other gas species ruins this tight correlation. The presence of a particular pattern in a spectrum thus leaves the atmospheric scientists with several possibilities for which combination of gases is responsible for it. To address this issue, scientists utilize a method called *retrieval*. In brief, retrieval involves calculating a variety of possible spectra using background knowledge about the distribution and concentration of atmospheric gases. Each of these possible spectra derives from a guess about what the actual gas concentrations are. Then, by using a

selection procedure that minimizes the difference between the real spectrum and the calculated one, taking into account relevant background knowledge, the scientists identify their best guess for the actual state of the atmosphere. Lusk argues that this is a case in which scientists explain data, and for good reasons. The spectrum is the data, and it is explained in the sense that a systematic derivation is provided for it. Lusk's argument for the systematic nature of the explanation relies on the fact that the atmospheric scientists effectively answer "what-if-things-had-been-different" questions about possible causal contributions to the spectrum. Answers to these counterfactual dependence questions are provided by the calculation of various plausible spectra, given the relevant background knowledge, which spectra are then used in the selection procedure to determine which atmospheric state to infer. Lusk suggests that not only is this case an example of the systematic explanation of data in scientific practice but also that the explanations in question play a crucial role in the identification of the phenomena of interest. Lusk also points out that the inverse problem of the sort involved in his atmospheric science case can be found in several other areas of science as well, such as geology, climatology, astronomy, medical imaging, and so on. Thus, he argues that the role of explaining data may be significantly more influential in scientific practice than Bogen and Woodward originally countenanced.

Noise

Bogen and Woodward distinguished the idiosyncratic causes contributing to data from the phenomenon of interest by describing phenomena as having stable and repeatable characteristics. In several papers, they characterize the idiosyncratic contributions to the causal production of data as *noise* or *error*. In light of this, one might wonder whether there are types of noise and/or error that are important to science in practice, which have stable and repeatable characteristics themselves. For instance, *shot noise* is a particular type of noise well-known to users of electronics. Photons or electrons, considered as discrete quanta, arrive at a detector shot by shot. Consider constructing an image using a CCD camera one photon at a time. The image would be *noisy*, and its clarity

would increase as more and more photons were added. Knowing that this sort of noise has this stable and repeatable characteristic—namely that it diminishes in a predictable and formalizable way—and that it has a theoretical explanation—in this case resting on quantum discreteness—is epistemically useful to scientists working with such detectors. This suggests that either there are some useful explanations of noise or that in some cases, the noise is (first) the phenomenon of interest.

Woodward (in his 2010 article for instance) emphasizes that scientists do not explain single data points. There may be some cases, however, in which a high-level explanation of even a single data point is both present and desirable in scientific practice. To take another example related to detector electronics—a spurious datum can be caused by a single high-energy cosmic ray passing through a detector. A high-level theoretical explanation of the influence of a cosmic ray on the production of data can be given, and the contribution of the *artifact* or *error* sufficiently characterized to warrant removing the affected datum from the main scientific analysis (unless the phenomenon under study is cosmic rays).

The Nature of Phenomena

Drawing a distinction between data and phenomena also raises questions about the nature of the phenomena. Following Bogen and Woodward, phenomena are supposed to be stable and repeatable across some range of contexts. They are immune to some idiosyncrasies that vary between such contexts, and they are not one-off events or flukes. But how do scientists identify phenomena or decide upon whether something counts as a phenomenon or not? David Colaço has argued that in addition to attending to explanations of phenomena, attention ought to be paid to how phenomena are characterized in the first place, and indeed how they are sometimes recharacterized in light of empirical research. Colaço's argument responds to the so-called mechanists in philosophy of science who have argued that phenomena are recharacterized in response to discrepant empirical results. For example, scientists may initially regard and represent what they ultimately deem to be multiple distinct phenomena under a single initial characterization. Colaço

contends that the mechanists mistakenly held that when characterization of a phenomenon faces discrepant results, scientists will recharacterize that phenomenon, for example, by splitting up what they regard as an inappropriately lumped characterization in response to the discrepant results. In contrast, Colaço argues that faced with such discrepancies, scientists suspend judgment regarding the initial characterization and pursue further research to investigate the discrepancies, rather than rejecting the characterization outright. Scientists do reject phenomena characterizations when faced with empirical results that are incompatible with those characterizations, such as when conditions supposed to produce the phenomenon of interest fail to do so.

Restatement and Defense

In 2011, Woodward explicitly revisited the arguments of Bogen and Woodward from 1988, providing clarifications and responding to some points of criticism that had been raised in the intervening years. Woodward clarified the scope of the argument about data and phenomena—rather than insisting on a sweeping generalization about what *always* or *typically* happens in scientific reasoning, he emphasized the desirability of paying attention to various kinds of reasoning from data-to-phenomena that are not as prominent in philosophy of science as one might expect given their centrality in scientific practice. In his reflections on the original argument, Woodward does still stand by the original position of Bogen and Woodward regarding explanations of data. In particular, Woodward continues to emphasize that scientists do not seek to provide such explanations both because the scientists do not need them and because such explanations would be impractical to furnish.

The original Bogen and Woodward article and the subsequent work by both of these scholars separately and in collaboration emphasized a second point about scientific practice. Whereas philosophers of science proceeding from the empiricist tradition have often highlighted the importance of observation and observations in science, Bogen and Woodward point out that many of the phenomena of scientific interest are not strictly *observable* in the sense of being perceivable by human senses. Moreover, in focusing philosophical

attention too narrowly on observations, philosophers of science have missed aspects of data collection and processing that are important for the epistemology of science.

Nora Mills Boyd

See also Big Data; Data Models; Empiricism; Evidence; Experiment, Theory of; Explanation; Fact Versus Theory; Statistics

Further Readings

- Bogen, J. (2002). Experiment and observation. In P. Machamer & M. Silberstein (Eds.), *The Blackwell guide to the philosophy of science* (pp. 128–148). Blackwell.
- Bogen, J., & Woodward, J. (1988). Saving the Phenomena. *The Philosophical Review*, 97(3), 303–352. <https://doi.org/10.2307/2185445>
- Colaço, D. (2020). Recharacterizing scientific phenomena. *European Journal for Philosophy of Science*, 10, 14. <https://doi.org/10.1007/s13194-020-0279-z>
- Leonelli, S. (2009). On the locality of data and claims about phenomena. *Philosophy of Science*, 76(5), 737–749. <https://doi.org/10.1086/605804>
- Lusk, G. (2021). Saving the data. *British Journal for the Philosophy of Science*, 72(1), 277–298. <https://doi.org/10.1093/bjps/axy072>
- Woodward, J. (1989). Data and phenomena. *Synthese*, 79(3), 393–472. <https://doi.org/10.1007/BF00869282>
- Woodward, J. (2010). Data, phenomena, signal, and noise. *Philosophy of Science*, 77(5), 792–803. <https://doi.org/10.1086/656554>
- Woodward, J. F. (2011). Data and phenomena: a restatement and defense. *Synthese*, 182, 165–179. <https://doi.org/10.1007/s11229-009-9618-5>

DATA MINING

See Big Data

DATA MODELS

Data models, or *models of data*, are processed versions of data, typically prepared in such a way as to make the data usable as evidence. The term *first*

gained prominence in a 1962 paper by the philosopher of science Patrick Suppes, and after a period of relative neglect has now become an area of active research. This entry summarizes some of the key distinctions and debates, including those regarding the ontology of data, the difference between raw data and data models, the purpose(s) of data modeling, the variety of data processing methods used in the production of data models, and how data models should be evaluated.

All STEM research—indeed all empirical inquiry—is heavily reliant on data collection, processing, and interpretation; hence data models are of foundational importance. Conceptualizing the research process as involving data modeling helps to clarify the relationship between data and the world, data and theories, and the role of modeling in the scientific process.

Ontology of Data and Data Models

One of the most basic theoretical questions one can ask about data concerns what kinds of things data are. Some researchers place no restrictions on which elements of the world can count as data. On this view, laboratory mice could themselves count as data. Other researchers define data as *records* of a process of inquiry involving a physical interaction between the researcher, measuring instrument, and the world: It is only the results of observations or measurements performed on the mice colony that count as data. On this latter view, data necessarily involve a level of abstraction. Data involve replacing the thing in the world with a number, a photograph, or a recorded description. Which view one adopts arguably has epistemic implications such as whether data are unproblematically *given* or whether they are subject to the sort of questions raised by the philosophy of observation, philosophy of experiment, and philosophy of metrology (measurement).

What kind of thing is a data model? When Suppes first introduced the notion of a data model, he conceptualized them as Tarskian or set-theoretic models. Although Tarskian models are prominent in mathematics, the more commonly used notion of model in science today treats models as *representations* of some target system. Some have argued that this representational notion of a data model is

important for recognizing that data are *about* the world, and hence for making sense of their ability to serve an evidential role.

Distinguishing Raw Data, Data Models, and Models

Data models are typically contrasted with *raw data*, but it is debatable whether and how such a distinction can be meaningfully drawn. Raw data are sometimes defined to be the immediate output of an observation, measuring process, or instrument. Often, however, such immediate outputs are not yet in a form that is usable by scientists. This can occur for a wide variety of reasons: First, the data-collection process might involve some obvious errors resulting in a certain portion of *bad* data that should be excluded from the data set. Second, there could be a source of *noise* in the data signal that needs to be filtered out or subtracted. Third, the output of the instrument might be of a quantity (measurand) that is closely related, but not identical, to the data quantity of interest, and hence must be converted before being useful. Fourth, data are often discrete samples of a continuous quantity and hence need to be smoothed or interpolated. Fifth, data often need to be organized or ordered before patterns can emerge. The outcome of these various processes of data wrangling—that is, of correcting, converting, smoothing, organizing, and so on—is what we call a *data model*.

The term *data model* can refer to any type of data product, including data sets, graphs, equations, images, and even processed artifacts. Statisticians often take *data model* or *model of the data* to refer more narrowly to equations that summarize a data set or the relationships therein. On this construal, models of data are akin to phenomenological models that summarize empirical relationships without representing or positing underlying mechanisms and processes. Phenomenological models, however, are typically distinguished from models of the data by playing a broader interpretive and inferential role. Data models can also be contrasted with data-driven models, which are typically constructed through machine learning and are trained on data sets to predict the value of a dependent variable from a set of independent variables.

On the other side, data models can also be contrasted with raw data. In practice, however, this distinction too can become blurred, with the term *raw data* often taking on a relative rather than absolute meaning. There are frequently many different layers of data processing, depending on the intended use of the data, not all of which are performed at the same time or by the same researchers. Scientists thus often use the term *raw data* to describe the data model at the stage before the data processing they are about to engage in. The raw data/data model distinction also becomes blurred as measuring instruments become more sophisticated, with a lot of the data processing (correction, conversion, etc.) happening within the instrument itself. Given the substantial, and often theory-mediated, processing involved, it is not clear that even the instrument output should be called *raw*. While some researchers are comfortable with the idea that it is “data models all (or almost all) the way down,” others have tried to recover a principled distinction, such as by taking raw data to be nonrepresentational and only becoming representational when they are processed as a data model.

The Purpose(s) of Data Models

When Suppes first introduced the concept of a model of the data, he conceived of it as part of a hierarchy of models, which included not only models of data but also models of experiment and models of theory. The purpose of a data model on this view was to provide a bridge for linking theories to data. Being tied specifically to the Tarskian notion of model (which Suppes used for both models of theory and models of data), the bridge was to be effected formally in terms of the presentation or instantiation of shared structure.

Most researchers, however, see the purpose of data modeling as helping data to function better as evidence. For those who adopt a representational view of models, this involves drawing out some signal or representational component of the data to make it a more accurate or transparent representation of that feature of the world. Here the aim of data modeling is not necessarily to serve as bridge to some theory but instead to

arrive at a better or more useful data product. Some researchers argue that, rather than being part of a synchronic hierarchy to theory, data modeling should instead be viewed as part of a diachronic data lineage. Although one may use a data model to test a theory, this is just one of many possible purposes.

How Data Models Are Produced?

Many different manipulations and processes can be involved in data modeling. Some involve statistical analysis methods such as data reduction, curve fitting, correlation analysis, regression analysis, statistical significance tests or calculation of confidence intervals, and error estimation. Other data modeling methods go beyond statistical analysis but are just as essential for making data usable as scientific evidence. Which techniques should be used in a particular case will depend on the nature of the data, the aims of the researchers, and the norms of their fields of study.

All data interpretation involves theoretical or conceptual assumptions about the experimental processes by which the data were produced, which are necessary for their proper semantic interpretation. Some data-processing methods, however, can make more substantial use of theoretical knowledge in the production of data models. One widely used example is data conversion, which involves converting a measure of one experimentally accessible quantity into data about another closely related quantity of interest. Examples of this include converting data about the travel time of some acoustic or electromagnetic signal into data about depth or distance, or converting data about the height of a mercury column in a thermometer into data about temperature. Another category of data-processing methods used in the production of data models falls under the rubric of data correction, which involves removing various sources of noise from the data or correcting for other errors, such as in a process of data calibration. This may be particularly important in cases where the data have been collected in the field rather than in a controlled lab setting. A third process is data interpolation, which involves filling gaps in the data, especially when the collected data represent some continuous quantity or need to be evenly distributed in space or time for a context of

use. These are just a few examples of the many ways raw data are turned into data models, and in many cases several of these data-processing methods will be used together.

For those who think of physical objects themselves (and not just numerical or other records of those objects) as data, the production of associated data models requires methods other than statistical manipulation, such as specimen preparation (e.g., preparing fossils) or image-processing techniques for photographs. Data modeling can also involve the preparation or ordering of data for storage and dissemination, including the transfer of data from one substrate or vehicle to another (such as in transferring data from descriptive text in a book to a computer database).

There remain substantive questions about how the lines between data modeling and various other data practices, such as *data curation* or data *visualization*, should be drawn. Indeed, standardizing the definitions of these and other terms related to data processing is an ongoing project.

Evaluating Data Models

When thinking about data as the output of a measurement, data are typically evaluated in terms of accuracy, precision, and resolution. These three terms are often explicated using the intuitive analogy of a dartboard: *accuracy* is how close the dart lands to the bullseye; *precision* is how closely together a number of darts cluster (anywhere on the dartboard), and *resolution* is how thin or thick the dart point is. Precision, however, is arguably better thought of as describing the conditions of measurement, rather than the quality of resulting data. More useful is the notion of measurement *uncertainty*, which can arise from many different sources and whose estimate should accompany the data, typically in the form of metadata (i.e., data about data). The quality of data thus can also depend on the quality of the metadata that accompanies the data model. There are substantive issues in both philosophy and metrology concerning exactly how these various concepts should be defined and deployed.

In addition to the epistemic evaluation of data models, involving for example how close data values are to some *true value*, there are socially driven criteria for evaluating data models, which

are often articulated in terms of the so-called FAIR data principles. FAIR is an acronym for Findability, Accessibility, Interoperability, and Reusability. The FAIR principles are taken to apply “not only to ‘data’ in the conventional sense, but also to the algorithms, tools and workflows that led to that data” (Wilkinson et al., 2016, p. 1). These principles are central to the *open data* movement.

More generally, scholars in both the philosophy of science and metrology communities have argued that data models should be evaluated, not *in abstracto*, but rather contextually in terms of a multidimensional problem space that assesses their adequacy or fitness for particular purposes. In the *adequacy-for-purpose* view of data models, the relevant question is not just how closely the data model resembles some feature of the world, but whether or not it is an adequate data model given a certain set of resources, methodologies, and other constraints for accomplishing a particular scientific aim. This view recognizes the essential role that context plays in data evaluation; a data model of a certain accuracy, and so on, that is adequate for a scientific purpose in one context may or may not be adequate in another context.

Yet another dimension along which data models can be evaluated is ethical. Data can be collected, curated, modeled, and disseminated in ways that are biased against—and harmful to—particular communities or social groups, such as women and Black, Indigenous, or other People of Color (BIPOC). Data models can embed or obscure systematic biases, especially in social contexts such as medicine, finance, and criminal justice. For example, the Black Lives Matter movement has called attention to the ways in which racial data collected by law enforcement can be biased along many dimensions, from over-policing certain social groups to biased police reports, leading to what AI Now researchers have called *dirty data*. These data ethics issues become especially prominent when data models are used for machine learning, artificial intelligence, and other automated decision making. Such data arguably require not just an evaluation of the epistemic risks but a moral evaluation of the ethical risks as well. The Global Indigenous Data Alliance has argued that the FAIR principles of data evaluation

are inadequate in that they ignore power differentials and historical contexts. They argue for a complementary set of what they call CARE principles for data evaluation, which stands for Collective benefit, Authority to control, Responsibility, and Ethics.

Alisa Bokulich and Aja Watkins

See also Accuracy; Big Data; Data and Phenomena; Evidence; Measurement; Modeling; Philosophy of Science; Statistics

Further Readings

- Bokulich, A. (2020). Towards a taxonomy of the model-ladenness of data. *Philosophy of Science*, 87(5). <https://doi.org/10.1086/710516>
- Bokulich, A., & Parker, W. (2021). Data models, representation, and adequacy-for-purpose. *European Journal for the Philosophy of Science*, 97(5). <https://doi.org/10.1007/s13194-020-00345-2>
- Harris, T. (2003). Data models and the acquisition and manipulation of data. *Philosophy of Science*, 70(5), 1508–1517. <https://doi.org/10.1086/377426>
- Humphreys, P. (2014). X-Ray data and empirical content. In P. Schroeder-Heister, W. Hodges, G. Heinzmann, & P. Bour (Eds.), *Logic, methodology and philosophy of science: Proceedings of the Fourteenth International Congress* (pp. 1–15).
- Leonelli, S. (2019). What distinguishes data from models? *European Journal for Philosophy of Science*, 9(2), 22. <https://doi.org/10.1007/s13194-018-0246-0>
- Leonelli, S., & Tempini, N. (Eds.). (2020). *Data journeys in the sciences*. Springer Open. <https://doi.org/10.1007/978-3-030-37177-7>
- Mayo, D. (1996). *Error and the growth of experimental knowledge*. University of Chicago Press.
- Suppes, P. (1962). Models of data. In P. Suppes (Ed.), *Studies in the methodology and foundations of science: Selected papers from 1951 to 1969* (pp. 24–35). Springer Netherlands. https://doi.org/10.1007/978-94-017-3173-7_2
- Wilkinson, M. D., Dumontier, M., Aalbersberg, I. J., Appleton, G., Axton, M., Baak, A., Blomberg, N., Boiten, J.-W., da Silva Santos, L. B., Bourne, P. E., Bouwman, J., Brookes, A. J., Clark, T., Crosas, M., Dillo, I., Dumon, O., Edmunds, S., Evelo, C. T., Finkers, R., . . . Mons, B. (2016). The FAIR Guiding principles for scientific data management and stewardship. *Scientific Data*, 3, 160018. <https://doi.org/10.1038/sdata.2016.18>

DEDUCTION

Deduction is one of the two chief approaches to human thinking, the other approach being induction. Although some philosophers have proposed other approaches such as abduction, the two approaches most widely recognized by logicians remain deduction and induction. Of those two, deduction has received relatively more attention and appreciation in the history of logic and theory development and constitutes the major form of modern logic.

Deduction Defined

According to Aristotle (385–322 B.C.E.), deduction is inference from the general to the particular. Although he discussed both deduction and induction in his book the *Organon* (*The Instruments*), his primary focus was deduction, mainly in the form of his syllogism, which is still deemed as representative of deductive logic. The *Oxford English Dictionary* adopts Aristotle's definition and defines deduction as "the process of deducing or drawing a conclusion from a principle already known or assumed; spec. in logic, inference by reasoning from generals to particulars; opposed to induction."

Later logicians have broadened the definition of deduction and put more types of inferencing methods into this category. Two widely used logic textbooks, Irving Copi and colleagues' *Introduction to Logic* (14th edition) and Patrick Hurley's *A Concise Introduction to Logic* (12th edition), both define a deductive argument as one whose premises support the conclusions in the way that it is impossible (as opposed to improbable in induction) for the premises to be true and the conclusions to be false. Hurley lists five types of typical deductive arguments: arguments based on mathematics, arguments from definitions, categorical syllogism, hypothetical syllogism, and disjunctive syllogism. Arguments based on mathematical calculations, measurements, and axioms (but not statistics, which forms the basis of inductive arguments, as they infer from a sample to the population) have the validity of their conclusion fully guaranteed from the premises. An argument from a definition is deductive because the information in the conclusion is fully contained in the definition, the premise.

A syllogism is a special format of deductive argument that has two premises and a conclusion. Categorical syllogism is the classic syllogism discussed by Aristotle, in which both premises and the conclusion are either universal statements or particular statements. For instance,

Major premise: All men are mortal.

Minor premise: Socrates is a man.

Conclusion: Therefore, Socrates is mortal.

Copi and colleagues and Hurley also classify the immediate inferences such as reduction and conversion of categorical propositions as deduction and develop other type of syllogisms in the deductive system, such as hypothetical syllogism and disjunctive syllogism. A *hypothetical syllogism* adopts the syllogism format but has a conditional statement as the first premise. For instance,

Major premise: If it rains, the floor will be wet.

Minor premise: It rains,

Conclusion: Therefore, the floor must be wet.

A *disjunctive syllogism* has a disjunctive statement as the first premise. For instance,

Major premise: The flower is either red or blue.

Minor premise: The flower is not blue.

Conclusion: Therefore, the flower must be red.

Philosopher and logician Karl Popper (1902–1994), who believed that deduction is the only logical approach to reasoning, provided a clear definition of deduction. That is, a deductive statement is true if and only if when all the truth contained in the conclusion has been transmitted from the premises. He further thought that the definition is a tautology to the following one: A deductive inference is valid if and only if no counterexample exists. That statement served as the philosophical foundation of his famous theory of falsification.

Early Deduction: The Aristotelian Syllogism

Aristotle's syllogism is the most influential early format of deductive reasoning. A *categorical syllogism*, which contains three propositions (two premises and one conclusion), states a relationship between two categories, the subject and the

predicate. Those propositions fall into one of the following four types: universal affirmative proposition A, universal negative proposition E, particular affirmative proposition I, and particular negative proposition O. Aristotle called the subject in the conclusion the *minor term*, as it normally stands for a smaller category in the conclusion. The predicate of the conclusion is called the *major term*, as it normally stands for a greater category. The premise that contains the minor term is therefore called the *minor premise*, and the premise containing the major term is called the *major premise*. The term that appears in the premises but not the conclusion is called the *middle term*. All Aristotelian syllogisms can be put in a standard form, in which all statements are in the standard form of A, E, I, or O and are listed in the order of major premise, minor premise, and the conclusion. When sorted into the standard form, the validity of the conclusions can be judged by merely examining the format of the syllogism: its mood and figure. *Mood* is determined by the type (A, E, I, or O) of the propositions, and *figure* is determined by the position of the middle term. In figure 1, the middle term appears as the subject of the major premise and the predicate of the minor premise; in figure 2, the predicates of both premises; in figure 3, the subjects of both premises; and in figure 4, the predicate of the major premise and the subject of the minor premise.

In figure 1, only the mood AAA, EAE, AII, and EIO are unconditionally valid; in figure 2, EAE, AEE, EIO, or AOO; in figure 3, IAI, AII, OAO, and EIO; and in figure 4, AEE, IAI, and EIO. In the Middle Ages, a mnemonic was created to help memorize the syllogism valid in each figure, which goes:

Barbara, Celarent, Darii, Ferio que prioris;
Cesare, Camestres, Festino, Baroko secundae;
Tertia, Darapti, Disamis, Datisi, Felapton,
Bokardo, Ferison, habet; quarta in super addit
Bramantip, Camenes, Dimaris, Fesapo, Fresison

The mood of the syllogism is indicated by the vowels in the words. The mnemonic also tells how to convert all moods in figures 2, 3, and 4 into a first-figure format. If a vowel is followed by an “s,” just switch the positions of the subject and the predicate of that statement represented by the

vowel; the mood will be converted into a first-figure mood. For instance, Cesare is a second-figure mood. The “s” after the first “e” tells us that we just need to switch the positions of the subject and the predicate in the major premise to convert Cesare to a first-figure syllogism, Celarent. If a vowel is followed by a “p,” the mood can be converted into a first-figure mood by switching the positions of the subject and the predicate of the statement associated with that vowel, and the statement needs to be limited in quantity by a quantifier (converted from a universal statement to a particular statement). For example, Darapti refers to AAI in the third figure. The “p” following the second “a” shows that an AAI syllogism should be converted to the first-figure mood AII through switching the positions of the subject and the predicate of the minor premise and reducing the minor premise from a universal statement into a particular statement. An “m” following a vowel indicates the need of switching the positions of the major premise and the minor premise. For instance, the third-figure IAI (Disamis) needs to switch its major premise and minor premise (also switching the positions of the subjects and the predicates of the original major premise and the conclusion as indicated by the two “s”s) to be converted to the first-figure mood AII.

The Aristotelian syllogism must follow six rules, any violation of which commits a fallacy:

1. Avoid four terms: a syllogism can only have three terms: the major term, the minor term, and the middle term (which by definition should connect the major premise and the minor premise but not be in the conclusion).
2. The middle term should be distributed (all elements in the category described by the term are covered) at least once. An example of the middle term not being distributed at least once is: “all Frenchmen are Europeans” and “all Germans are Europeans,” from which one cannot conclude that “all Frenchmen are Germans” or any other thing.
3. If the minor term or the major term is distributed in the conclusion, it also needs to be distributed in the premise.
4. Avoid two negative premises. At least one premise must be affirmative.
5. If either premise is negative, the conclusion must be negative. If both are affirmative, the conclusion must also be affirmative.

6. At least one premise must be universal. If one premise is particular, the conclusion must also be particular. If both premises are universal, the conclusion should also be universal.

In the Aristotelian syllogism, the universal propositions imply the particular propositions with the same subjects and predicates. If A is true, I must be true. Vice versa, if E is true, O must be true. For instance, if the universal proposition “all men are mortal” is true, then the particular proposition “some men are mortal” must also be true. On the other hand, the truth of the universal negative proposition “no men are immortal” implies the truth of the particular negative proposition “some men are not immortal.” British mathematician and logician George Boole (1815–1864), however, argued that although the universal statements A and E do not imply the existence of the objects described in the subject or the predicate terms, the particular statements I and O do imply such existence, meaning “at least one object existing” in the subject or predicate terms (particular statements are also called *existential statements*). So Boolean logicians acknowledge the validity of such universal statements as “all perpetual motion machines are a good thing” or “all perpetual motion machines need no external sources of energy.” But they believe that the particular (existential) statements such as “some perpetual motion machines are a good thing” or “some perpetual motion machines need no external sources of energy” are invalid, because those existential propositions’ assumption of at least one perpetual motion machine is not met in the world.

The Aristotelian syllogism dominated the classical Greek era and the medieval era, influencing Europe significantly. According to Thomas Kuhn in his influential book *The Structure of Scientific Revolutions*, some scientists believed that if it weren’t for the domination of deductive logic, some scientific revolutions might have come earlier.

Contemporary Deduction: Logic, Math, Computer Science, and Artificial Intelligence

The greatest contribution of Boole to logic is probably his origination of the algebra of logic. A revolution to the Aristotelian syllogisms that rely on figures and moods, Boolean logic builds an

algorithm system that can be applied to an unlimited quantity of argumentations, essentially deductive, and denote the underlying structure of logic. In his book *The Mathematical Analysis of Logic* (published in 1847), Aristotelian syllogisms are algorithmized into equations to calculate the logical propositions with the help of algebraic techniques. For instance, the proposition “All X is Y” is algorithmized as “ $x = xy$ ”; “No X is Y” as $xy = 0$; “Some X is Y,” as $v = xy$ (v stands for *some*); and “Some X is not Y,” or $v = x(1 - y)$. He later developed his algebra of logic into a binary system that contains only true or false values {T, F} in his book *The Laws of Thought* (published in 1854), which laid down the foundation for deductive logic to be introduced into computer science and electronic engineering research and turns into a core part of them. His algebra of logic was inherited and developed by William Jevons, Charles Sanders Peirce, Ernst Schröder, Edward Huntington, Marshall Stone, Thoralf Skolem, and eventually Alfred Tarski, who developed a system of equations to calculate the logical relations and establish the algebra of logic within the modern logical system.

Gottlob Frege (1848–1925) established a tradition that is superficially similar to, but significantly different in nature from, the Boolean tradition of solving logic problems with algebraic techniques. Frege created a system of arithmetic notations for formal logic to fulfill Gottfried Leibniz’s idea of *universal characteristic*, a language of pure thought. When Boole’s follower in Germany, Schröder, criticized Frege for merely mimicking Boole’s work without mentioning and advancing further from it, Frege replied and argued that what Boole had created was merely Leibniz’s *calculus ratiocinator*, a system of uninterpreted markers for calculating without thinking. In his book *Begriffsschrift* (Concept-writing, published in 1879), Frege clarified that he was not creating an artificial similarity that makes concepts as a sum of characteristic marks, which he implied to be the nature of the Boolean systems. Frege’s work was actually to reduce mathematics to logic in an arithmetical form, which can help prove all mathematical problems. In Frege’s tradition, which has since been inherited by most logicians and mathematicians, mathematics is a special form of logic.

Frege’s system, however, was discovered by Bertrand Russell to have a contradiction, which later became known as *Russell’s paradox*: The

systems contains a paradoxical object: the set of all sets that are not members of themselves. To resolve the paradox, Russell introduced the notion of *logical type* and developed the theory of types, which is thoroughly discussed in the book he coauthored with Alfred North Whitehead, *Principia Mathematica*, the most popular work of logic in the 20th century. The discussion of Russell's paradox also inspired the exploration of the completeness of logic systems; namely, whether the elements of a logic system can all be deduced from its axioms. In 1931, Kurt Gödel published a paper in which he proved that no mathematical systems sufficient to embed arithmetic are deductively complete; that is, all logical or mathematical systems contain statements that can be neither proved nor disproved. Further, those systems cannot even prove the consistency of themselves. Known as *Gödel's incompleteness theorems*, those findings have caused many to suspect that a complete account of deduction is not yet achieved and probably will never be.

Frege also initiated the thread of research on predicate analysis, which, unlike traditional propositional analysis, uses concepts rather than sentences as the unit of logical analysis and therefore allows more delicate and sophisticated calculations and inferences, such as quantification analyses. The establishment of predicate logic lays down the foundation of modern logic. In Frege's predicate logic system, the most well-known symbols are the existential and universal quantifiers ($\exists x$ and $\forall y$). Frege's focus was on first-order predicate logic, in which quantifiers are applied only to variables that stand for individual objects, although he also covered second-order predicate logic, in which quantifiers are also applied to the sets of individual objects. Second-order predicate logic and even higher order predicate logic, in which multiple layers of sets of individual objects are quantified, were further developed by other logicians and eventually in *Principia Mathematica*, enabling logical analyses of arbitrarily complex situations.

The Deductive Scientific Inquiry Models

Inductive logicians have long believed that deductive logic is useful in preparing arguments but useless in scientific inquiry, because it generates no

new knowledge. On their view, induction is the only way to help people reason about phenomena outside the existing sphere of experiences and gain new knowledge.

Deductive logicians disagree, in a debate going back to ancient Greece between the Stoics and the Epicurean philosophers. Ever since the Scottish philosopher David Hume (1711–1776) pointed out the problem of induction, which reasons from the past to the future, philosophers have tried to solve the problem and develop deductive models of scientific explanation. Some philosophers, such as Popper, totally rejected induction and marked deduction as the only valid way of thinking, reasoning, and scientific inquiry. Among those deductive models, Karl Popper's falsification model, Carl Hempel's deductive-nomological (D-N) model, and the hypothetico-deductive (H-D) model are the most well-known.

Popper's Falsification Model

Popper held that it was impossible to objectively draw patterns from single facts, because all facts, when they were described, were organized on the basis of certain theories that the scientists already held. Although often perceived as a member of the logical positivism movement, Karl Popper was actually a critic of that movement, who maintained that all scientific inquiry was conducted with a system of theoretical assumptions and therefore could not start from scattered experiences and observations, as claimed by the logical positivists.

Believing that theories, or any general statements, could be logically never proved but only disproved when even one piece of counterevidence appeared, Popper argued that falsifiability was the key difference between science and pseudoscience, and that falsification was the only way to test theories. Karl Marx's theory of history, Sigmund Freud's psychoanalysis, and Alfred Adler's individual psychology were all considered as pseudoscience by Popper, because they could be used to explain any situation and no case could be used against those theories. They also could not make precise predictions. Albert Einstein's theory of relativity was classified as science by Popper, because it could develop predictions that can be tested. Such testing, however, can never prove a

real science, even if the evidence from the tests supports the scientific theory, such as in the case of the relativity theory tests. Scientific theories could only be falsified when the evidence is against them. For any real scientific theory constructed through deductive reasoning, Popper said, scientists could search for an opposite case. If the search was successful, the deduction was invalid and the theory was rejected; otherwise the hypotheses survived the falsification. Whereas inductivism scholars insist that scientific knowledge is accumulated through direct observation, Popperian deductivism scholars hold that scientists cannot generate hypotheses through direct observation, nor can they prove the hypotheses through it. Scientists can only build hypotheses through thinking and then attempt to falsify the hypotheses through direct observation. Hypotheses that can explain many phenomena and withstand multiple attempts at falsification should be deemed as good theories.

But Popper's simple falsification method could easily reject an established theory with just one demonstrable instance of counterevidence. His student and follower Imre Lakatos therefore proposed a methodology of scientific research programs (MSRPs) to improve the tolerance of the falsification method. Lakatos's MSRP classifies a theory's contents as the central theses (hard core) and the peripheral theses (auxiliary hypotheses). Disproving the auxiliary hypotheses does not refute the theory; it actually improves the theory by enhancing its predicting capabilities. A theory is only falsified when the hard core is disproved. Meanwhile, Lakatos, unlike Popper (whose subject of examination was normally a single theory), liked to think of theories as elements in a sequence that shared the same hard core. His revision to Popper's falsification method allows many useful theories to pass falsification testing. Even so, in the domain of social or natural scientific research, the falsification approach has rarely been used.

The Deductive-Nomological Model

Carl Hempel's (1905–1997) D-N model is a well-known model based on John Stuart Mill's idea that the only type of scientific explanation has the *explicandum* (explanandum), the thing to be explained, as a logical consequence of the

explicans (also called *explanans*), the thing that explains. The D-N model contains three elements: (1) A lawlike statement or theory as the premise, stating a causal relationship between factor *x* and factor *y*, such as “whenever *x* happens, *y* happens”; (2) a statement of an antecedent condition of the happening of *x*; and (3) a deductive conclusion about the consequential happening of *y*. The model is called *deductive* because it requires the explicandum to be the logical consequence of the explicans; it is called *nomological* because it requires at least one *law of nature*, a universal law that describes a class of phenomena, to be the premise. It is also referred as the *covering law model* in that sense. A simple example of the D-N model of explanation would be:

Theory (lawlike statement): Increase of the volume of media coverage on an issue makes people think of that issue as more important.

Antecedent: The volume of media coverage on education has increased.

Conclusion: People will think of education as more important.

In scientific explanation practice, the theory and the antecedent each can be a set of statements. When the theory or the lawlike statement(s) is replaced by one or more statistical statements, the D-N model becomes an inductive-statistical model, a model popular among modern inductive scientists.

Hempel's D-N model has been criticized because it assumes a universal law whose truth has been established. This type of universal law, some philosophers point out, is impossible to be found and proved. Also, no evidence can support a universal law to be true in the future, as Hume argued in his statement of the problems of induction. British philosopher Stephen Toulmin used Charles Darwin's theory of evolution as an extreme example to show that the D-N model did not explain all the theories, as there was no universal law to guide Darwin's data collection and his articulation of the theory. Toulmin and other philosophers who follow Ludwig Wittgenstein's teaching reject the idea that universal laws were the prerequisite to explain human behaviors. Thomas Kuhn also pointed out that historical studies were not based on lawlike statements. The

D-N model, marketing researcher John O'Shaughnessy maintains, also cannot explain the causal relationship that many scientists are seeking, because the D-N model provides necessary but not sufficient conditions for causal explanation. Even that necessary causal relationship is not fully established, because in the D-N model only the relationship and time order are examined, and no effort has been made to exclude the third factor. The consequential changes of x and y , therefore, can be both caused by a third party, z . Some scholars believe that the D-N model could be a great model of scientific explanation if the *why*, which addresses the underlying mechanism of causation, is included. Political scientist John Gunnell argues that although Hempel's D-N model has dominated natural science and, subsequently, social science, it is actually not an appropriate methodological paradigm for social science. He further notes that the deductive model is deficient even for natural science. According to Gunnell, Einstein, whom deductive logician Karl Popper often cited as a good example of deductive theorists, had never claimed that the relativity theory was generated through a deductive approach of thinking.

The Hypothetico-Deductive Model

The H-D model, which Popper also called *deductivism*, emerged in physics studies in the 17th century and became a major model guiding later natural and social sciences. It contains 11 steps, as summarized by O'Shaughnessy. In the first step, the scientists start to be interested in an issue and collect some perceptual experiences. In the second step, those perceptual experiences are sorted out under the guidance of existing knowledge. In the third step, when there are some difficulties to sort out those perceptual experiences into the existing knowledge system, research problems emerge. In the fourth step, to solve those problems, scientists generate speculative hypotheses. In the fifth step, a set of those related hypotheses, based on existing knowledge consistent to the hypotheses, are formulated into a tentative theory or model. Then in the sixth step, the scientists design a procedure (i.e., an experiment) to test the tentative theory. In the seventh step, data

are collected through the procedure to test the hypotheses. The eighth step, verification, shows whether or not the statistical tests in the procedure rise above the significant standard. If the hypotheses withstand the testing (the ninth step), they then are elevated as laws or formal theories in the tenth step and further become the foundation to explain other phenomena in the eleventh, and final, step.

In addition, O'Shaughnessy divides the H-D model into six stages: domain image, problem identification, speculative hypothesis, tentative model, prepare for testing, and testing.

1. *Domain image.* In each field, a system of presuppositions exists to define the boundary of the academic domain. Actually, even scholars in the same domain may have different biases in their theoretical or methodological backgrounds. Those domain images or researchers' own backgrounds may not be right, but they are the fundamentals on which those researchers construct their theories.

2. *Problem identification.* All research projects start with goals of solving certain problems. Those problems are inconsistencies or puzzles that the researchers have identified in the conceptual system of their domains. Although solutions to those questions can be from all domains, researchers tend to find the solutions from their own fields, owing to their theoretical and methodological bias.

3. *Hypothesis.* The speculative solutions to the problems that researchers come up with are hypotheses. Hypotheses to scientists, as Popper stated in his autobiography *Unended Quest*, are like nets to fishermen. Scientists who do not use hypotheses cannot capture knowledge. With the increasing emphasis on the function of hypotheses, social scientists have also been studying how to effectively generate hypotheses. Social psychologist William McGuire, for instance, recommended about 100 techniques to help researchers generate hypotheses, although most of the methods are inductive in nature.

4. *Tentative model.* When the hypotheses are connected with each other and with existing concepts and propositions in the knowledge system

in that domain, a tentative model is formulated. Tentative models are postulated, rather than evolved from induction. Based on that statement, H-D scholars also believe that science is more like being invented rather than being discovered. It is essentially a perceptual reconstruction of the object. It describes the relationships among the concepts related to the problem under study. When the underlying mechanism of the relationships is included in the model to explain why those relationships exist, the model is further called a *theory*, a term that contains more substance in the philosophy of science.

5. *Preparing to test.* Preparation for testing is basically a process of connecting the theoretical concepts to the observable world through conceptual analysis. The first step of conceptual analysis, as communication scholar Steven Chaffee points out, is meaning analysis, through which a concept's intention (connotation) and extension (denotation) are articulated. Based on meaning analysis, definitions should be provided. Although analytical definitions, which state the relationships of concepts to their superordinate concepts, have been the traditional format of definition, operational definitions are instrumentally more important format of definition that theorists need. Operational definitions connect the concepts to observable indicators that scientists can measure with human senses, with or without the help of research devices. The research on how to develop operational definitions for theoretical concepts, also called *constructs* in some textbooks, is known as *measurement*. Operational definitions must be constructed based on thorough understanding of the analytical definitions. Certain concepts need no operational definition, because they are not abstract or far away from people's senses, such as people's demographical factors, or because those concepts are created based on empirical measures, such as economic indices.

6. *Testing hypotheses.* Traditionally, scientists test a hypothesis by first deducing an observable logical consequence from it and then testing the logical consequence. If the consequence is tested and observed, the likelihood of the hypothesis being true is raised. If the confirmation fails, the

hypothesis is not supported, but need not necessarily to be abandoned.

The influence of the H-D model can be seen from a model of scientific inquiry in a book written by education researcher Frank Kerlinger, *Foundations of Behavioral Research*. Kerlinger, on the basis of John Dewey's analysis of reflective thinking, proposes a four-step scientific approach, which has become influential among not only education researchers but also researchers in many other social science fields, such as communication, marketing, and psychology. In the first step, problem-obstacle-idea, scientists experience some difficulties in explaining certain things and formulate a statement of the problem. In the second step, based on past and/or present experiences, the scientists turn the problem into a hypothesis, a conjectural or tentative proposition. In the third step, the scientists conduct reasoning-deductions. This, according to Kerlinger, is the most important yet often overlooked step. In order to further test the hypothesis, the scientists need first to deduce testable consequences from the hypothesis. In the fourth step, the scientists find ways to test the consequences deduced from the hypothesis. The process is not necessarily always in that chronological order, and latter steps may happen before the previous steps have been completed. Preconceived ideas (problems or hypotheses) are the guides for collecting relevant facts, testing the hypotheses, and accumulating knowledge. Kerlinger highly values the role of hypothesis in his model. Hypotheses, he wrote, are working instruments of theory, without which no science would be possible, because hypotheses can be tested and shown as correct or incorrect apart from their creators' opinion or evaluation. Whether or not the hypotheses are confirmed, knowledge is advanced. The Kerlinger model of scientific inquiry is essentially a miniature of the H-D model.

Despite its domination in academia even today, deductive models of theory development, in particular the H-D model, cannot explain all the significant discoveries in many fields. In addition to Toulmin's citing the evolution theory as a counterexample, O'Shaughnessy also cites voyages of discovery as a good counterexample of the

deductive model of scientific exploration. An instance of scientific inquiry that arrives at a wrong conclusion is not totally useless. Voyagers often fail to find what they seek, while finding other things of great value and interest.

Qingjiang Yao

See also Completeness; Conceptual Analysis; Hypothetico-Deductivism; Induction

Further Readings

- Chaffee, S. (1991). *Explication*. Sage Publications.
- Copi, I. M., Cohen, C., & McMahon, K. D. (2010). *Introduction to logic* (14th ed.). Prentice Hall.
- Fisher, R. (1955). Statistical methods and scientific induction. *Journal of the Royal Statistical Society. Series B (Methodological)*, 17(1), 69–78.
- Goldfarb, W. (2003). *Deductive logic*. Hackett.
- Hurley, P., & Watson, L. (2018). *A concise introduction to logic* (13th ed.). Cengage Learning.
- Kerlinger, F., & Lee, H. B. (1999). *Foundations of behavioral research* (4th ed.). Wadsworth.
- Kuhn, T. S. (2013). *The structure of scientific revolutions* (4th ed.). University of Chicago Press.
- O'Shaughnessy, J. (1992). *Explaining buyer behavior: Central concepts and philosophy of science issues*. Oxford University Press.
- Popper, K. (1982). *Unended quest: An intellectual autobiography*. Open Court.

DEVELOPMENTAL SYSTEMS THEORY

Developmental systems theory (DST) is a comprehensive theoretical framework on biological development, heredity, and evolution. A metatheory about complex systems in nature, DST spans the fields of philosophy, biology, and psychology, providing a structure for the study of interactions among multiple factors that drive development. It offers a thorough understanding of human development by combining explanatory forces of nature and nurture, genes and environment, and biology and culture. This entry explores the history of DST, its definitions and characteristics, key concepts and themes, and future directions.

History of DST

DST is a general theoretical perspective based on the rich history of work by scholars in the implementation and interpretation of developmental science. Unlike many other theories, DST is not an individual theory offering predictions for empirical testing and based on the work of one person. Rather, on the basis of discussions and research by numerous philosophers and biologists of the 19th and the 20th centuries, psychologists in the 1970s began to develop a systems approach to human development, grounded in developmental and behavioral psychology, which has attracted the interest of, among others, molecular biologists and philosophers of biology.

Philosophers and scientists have been interested in development of humans and other organisms since ancient times. As a result of the progress in anatomical studies in the 17th century, two chief viewpoints emerged. *Preformationists* asserted that the development of the organism was simply an unfolding of inherited structures. *Epigeneticists*, however, thought that the organism became more complex during development, in which physical structures develop from indistinct germinal material. By the 19th century, the continuing improvement of early microscopes enabled the analysis of egg cytoplasm and discussion of the processes of differentiation and maturation. Similar to the debates between nativists and empiricists in philosophy, these two camps also focused on innate complexity versus gradual development. These parallel discussions paved the way for the emergence of psychology and evolutionary biology.

In the late 19th century, instincts started to be studied in humans owing to the influence of Darwinian evolution on psychology. In studies of comparative psychology, nonhuman animals were experimentally tested in laboratories for their learning behavior in isolation from their instincts as observed in natural settings. Functionalist psychology in the early 20th century in the United States made extensive efforts to understand the role of instincts in human behavior. Psychologists such as John B. Watson (1879–1958) and his followers then turned away from instinct because of its innate and indefinite

nature. Behaviorism, according to Watson, was the science of observable behavior. Behaviorists' emphasis on learning and experience and devaluation of innate and heredity factors turned psychology into a nondevelopmentally oriented discipline.

Early studies of child development by G. Stanley Hall (1846–1924) and Arnold Gesell (1880–1961) were initially not accepted into mainstream American psychology because of their methodological and theoretical perspective. In Europe, ethological studies on animal learning and development, especially by Konrad Lorenz (1903–1989), turned the focus on instincts again. Instincts were essential for animal behavior and independent from, and in some ways superior to, learning experiences. Phylogenetic information providing a base for all innate behavior offered a genetic blueprint for development, whereas environment played a largely supportive role. Criticisms of Lorenz's work by the naturalist and animal psychologist Daniel S. Lehrman (1919–1972) rejected the dichotomy of innate instincts and learned experiences. Lehrman's integrationist perspective is accepted as one of the first discussions of DST.

The discovery of the structure and function of deoxyribonucleic acid (DNA) in the early 1950s and later studies shed light on the molecular-level processes of inheritance, development, and behavior. Subsequently, the notion of genes as the chief or sole determinant of behavior was replaced with a two-way perspective that recognized the role of the environment, experience, and learning as well as heredity. According to Gilbert Gottlieb's model of probabilistic epigenesis, proposed in 1991, genes identify the neural structure, which influences and is influenced by the physiological function, which also has a reciprocal relationship with behavior. In their 1992 book, *Developmental Systems Theory: An Integrative Approach*, Donald H. Ford and Richard M. Lerner first presented the DST. Building on a rich history of scholarly work, they emphasized contextualism and dynamic interactionism in the process of development. Since then, most research grounded in the DST perspective has focused on human behavioral development, while DST-based philosophical discussions have contributed to the

theoretical discourse of molecular and developmental biology.

Definitions and Characteristics

Richard M. Lerner has outlined the defining features of developmental systems theories. According to Lerner, DSTs are relational metamodels that bridge many dichotomies such as nature–nurture, genes–environment, continuity–discontinuity, stability–instability, and biology–culture. All levels of organization within developmental ecology, from biology and psychology to culture and history, are integrated. Development is regulated by mutually influential connections between all levels of the system such as family and community, or individual and social context. The basic unit of analysis in human development consists of integrated actions between person and environment. The temporality and plasticity of human development allows different trajectories of intraindividual change. Plasticity of development is relative across the life span, the history of the individual, and the social/environmental context. Different levels and variables produce diversity and change between and within individuals, which is fundamentally significant for the study of human development. Applications of developmental science through interventions and policies can promote positive human growth. The integrated and multilayered nature of developmental systems requires scholars and researchers to exercise multidisciplinary knowledge and methodologies sensitive to change at all levels.

Within its multidisciplinary framework, DST is subject to certain inaccurate representations. First, DST is not just a philosophical perspective on biology but is based on, and in turn influences, experimental and empirical studies of development. Although DST is sometimes believed to emphasize only the developmental aspect of evolution at the expense of natural selection, in fact DST emphasizes the generative aspect of evolutionary development as well as its selective and adaptive nature. Some others have assumed that DST downplays the role of genes in development. On the contrary, DST recognizes the effects of genes as well as nongenetic factors on development and includes both factors in developmental

systems. The DST perspective posits that the genes do not simply code or specify behavior but, rather, influence the way a particular behavior develops. Roles that genes play in development are neither minor nor simple, but they are not in competition with nongenetic factors; rather, genes are in coevolution with them.

DST is also different from many other fields of development such as quantitative behavioral genetics and nativist cognitive developmental psychology. It even developed as a response to the formulations and methods of these fields. Behavioral genetics tried to account for the relative contributions of genes and environment to individual differences. DST objected to the focus on relative weights of factors and to the supposed duality of nature and nurture. Instead, it aimed at a process explanation of developmental mechanisms and their dynamic interactions. Nativist cognitive developmental psychology, exemplified by Noam Chomsky's linguistic acquisition device, posits innate faculties in several neurocognitive domains that shape children's cognitive development. DST also criticizes the nativist approach for its predetermined maturational view on development that minimizes the power of environmental forces.

Key Concepts and Themes

In the context of DST, epigenesis and developmental dynamics stand out as two core concepts. In contrast to preformationism and predeterminism, epigenesis puts forward that development unfolds by leading to increased complexity. It is not simply formed or determined in advance by biological structures or environmental factors. Similarly, epigenetics focuses on mechanisms and processes of genotype leading to a certain phenotype. *Epigenotype* includes multiple genes in the expression of a phenotype. DST enriches epigenotype by including nongenetic factors among the causal factors of the developmental process. Owing to the probabilistic nature and plasticity of development, a certain phenotype cannot be caused or determined by the mere addition of these factors. Rather, development is a contextual, interactive, and dynamic process.

The concept of developmental dynamics implies that the developing organism is in constant

dynamic interaction with its environment during which both the organism and the environment change. There are bidirectional influences between genetic and neural systems as well as between the behavior and the environment of the organism. Similar to the complex system of biological structures, environment is composed not only of physical but also social and cultural environment. For an accurate representation of development, these dynamic interactions between multiple layers should be studied over time in the course of development.

In their seminal work, Susan Oyama, Paul E. Griffiths, and Russell D. Gray outlined six key themes of DST as (1) joint determination by multiple causes, (2) context sensitivity and contingency, (3) extended inheritance, (4) development as construction, (5) distributed control, and (6) evolution as construction. According to the first theme, development is shaped by multiple factors that work together in an interactive fashion rather than an additive style. As for the second, development takes place in a multilayer context, and the context determines which factors will dominate others during this process. Third, the development and life cycle of the organism is driven by a broad range of inherited genetic resources. In the fourth theme, development is a constructive process that is not solely determined by preexisting inherited traits but constructed by the developing organism throughout the process. The fifth theme implies that development is not controlled by one type of factor alone but shaped by a fluid and diffuse set of factors. In the sixth theme, development includes evolution of the organism–environment systems rather than the evolution only of individuals or populations.

DST explains development through dynamic and mechanistic explanations: mechanistic because they present arrangements of factors within developmental systems, which lead to those processes; dynamic because processes result not only from the arrangements of factors but also from the dynamic operation of the mechanism. Bidirectional influences between factors from different levels work via top-down or bottom-up causal mechanisms. Such factors do not operate in isolation but work in a context of a whole set of other variables. In truly developmental explanations, the aim is not to explain causes of individual

differences but to explain developmental systems' potential to lead to these differences and other consequences.

Future Directions

Recent writings on DST have tried to offer integrationist views together with other perspectives such as evolutionary developmental biology (EDB) and evolutionary psychology as well as the Advaita philosophy of nonduality. Different interpretations of inheritance and an emphasis on genes led to a mutual exclusion between DST and EDB. A systematic and cross-disciplinary research program may help to close the gap between the two. Both DST and evolutionary psychology make use of self-organized patterns of development from genes to culture as well as probabilistic mechanisms of development in ontogeny (life span development) and phylogeny (across generations). A dynamic optimization method has been found to be useful for the integration of the two fields. Advaita philosophy, originating in the Vedanta tradition of India, and similar to the nondualistic aspect of DST, proposes a holistic and interconnected view of individual and cosmos.

In addition to the theoretical expansions, there have been conceptual and methodological advances offering clues about the future directions of DST. Integration of key concepts such as probabilistic epigenesis, dynamic interaction, and self-organization has led to a relational DST. The relationist view provides a richer understanding of development with embodied actions of a holistically organized developmental system. The relational biopsychosocial model synthesizes person, biology, and the sociocultural, whereby each component develops through co-construction and coevolution. Furthermore, the introduction of simulation models and dynamic modeling techniques has opened new fronts to DST. Connections and directionality among variables at inter- and intraindividual levels are better represented through these methods. Advanced analytical tools provide better means of studying the nonlinear, self-organizing, and complex nature of development.

Asil Ali Özdoğru

See also Biology, Evolutionary; Complex Systems; Evolutionary Psychology; Systems Science

Further Readings

- Ford, D. H., & Lerner, R. M. (1992). *Developmental systems theory: An integrative approach*. Sage.
- Frankenhuis, W. E., Panchanathan, K., & Barrett, H. C. (2013). Bridging developmental systems theory and evolutionary psychology using dynamic optimization. *Developmental Science*, 16(4), 584–598. <https://doi.org/10.1111/desc.12053>
- Griffiths, P. E., & Gray, R. D. (2005). Discussion: Three ways to misunderstand developmental systems theory. *Biology and Philosophy*, 20(2–3), 417–425. <https://doi.org/10.1007/s10539-004-0758-1>
- Johnston, T. D. (2010). Developmental systems theory. In M. S. Blumberg, J. H. Freeman, & S. R. Robinson (Eds.), *Oxford handbook of developmental behavioral neuroscience* (pp. 12–29). Oxford University Press.
- Lerner, R. M. (2006). Developmental science, developmental systems, and contemporary theories of human development. In R. M. Lerner & W. Damon (Eds.), *Handbook of child psychology. Vol. 1: Theoretical models of human development* (6th ed., pp. 1–17). Wiley.
- Molenaar, P. C. M., Lerner, R. M., & Newell, K. M. (Eds.). (2013). *Handbook of developmental systems theory and methodology*. Guilford Press.
- Overton, W. F. (2013). Relationism and relational developmental systems: A paradigm for developmental science in the post-Cartesian era. In R. M. Lerner & J. B. Benson (Eds.), *Advances in child development and behavior. Vol. 44: Embodiment and epigenesis: Theoretical and methodological issues in understanding the role of biology within the relational developmental system, Part A: Philosophical, theoretical, and biological dimensions* (pp. 24–64). Elsevier. <https://doi.org/10.1016/B978-0-12-397947-6.00002-7>
- Oyama, S., Griffiths, P. E., & Gray, R. D. (Eds.). (2001). *Cycles of contingency: Developmental systems and evolution*. MIT Press.
- Ploeger, A. (2010). Evolutionary psychology, developmental systems theory and Advaita philosophy as metatheories: Are the three compatible? *Integral Review*, 6(3), 201–206.
- Robert, J. S., Hall, B. K., & Olson, W. M. (2001). Bridging the gap between developmental systems theory and evolutionary developmental biology. *BioEssays*, 23(10), 954–962. <https://doi.org/10.1002/bies.1136>

E

ECOLOGY

See Environmental Studies; Natural Resources

EMPIRICISM

See Epistemology; Evidence; Experiment, Theory of

ENGINEERING, CONTEMPORARY

While engineering includes a subdiscipline called *engineering science* (statics, dynamics, thermodynamics, etc.), engineering, even engineering science, is not a science strictly speaking, not even a practical science. No part of engineering—design, testing, operations, maintenance, documentation, disposal, or the like—seeks to understand any part of the world as such (as science typically does). Instead, engineering attempts to change the world. Understanding (when it occurs in engineering) is a mere byproduct of that attempt. The primary vocabulary of engineering consists of directions: designs, specifications, maintenance schedules, and so on. The search for ways to change the world is what makes engineering a technology—like architecture, accounting, public health, and so on. What makes engineering, including engineering

science, different from most, perhaps all, other technologies is that engineering attempts to improve humanity's material condition without any upper limit to the improvement. Other disciplines recognize an upper limit after which there is nothing more to do, for example, the public welfare (the upper limit of public administration) or health (the upper limit of medicine).

Engineering's emphasis on unlimited improvement makes *theory* in engineering different in important ways from theory in a science (strictly speaking). For the sciences, theory is an aid to understanding; for engineering, theory is a guide to action, not simply in research (as in science strictly speaking) but also in invention, manufacture, quality control, construction, maintenance, and disposal in the world. For engineers, understanding is itself only a possible means of improving the material condition of humanity. Consider, for example, how differently research into metal fractures might be conducted in science and in engineering. Scientists (typically, physicists) will want to develop a theory that identifies the forces contributing to a fracture, explain fracture in terms of those forces, and perhaps use that explanation to predict fractures. They will not be concerned with design, that is, with means of reducing fractures in a mixed material in ordinary use (such as the alloys in the wing of an airplane). Engineers will.

Theory in General

Whatever else it is, a theory is like a seat in a theater. It offers a particular perspective on the world understood according to a certain discipline.

A discipline consists of a limited vocabulary in which to express claims, *prima facie* modes of proof, or other formalized ways of acting. For example, physics tries to understand the world (*the physical world*) in terms of waves, mass, time, space, and other physical properties expressed mathematically and tested by observation, prediction, deduction, and experiment. Though physics has something to say about every particle in the universe, however large or small, it is not a *theory of everything*. It is, for example, silent about how to arrange the world to increase accessibility, safety, or comfort. Neither *accessibility*, *safety*, nor *comfort* is in its vocabulary. Those three terms belong to other disciplines, such as interior design or public health.

Engineering theory typically differs from theory in the sciences in focusing on usefulness rather than on understanding. Engineering theory typically differs from theory in other technologies in having as its purpose improving the material condition of humanity. Most technological disciplines serving the material condition of humanity (architecture, public health, interior design, etc.) attempt to be useful but not (like engineering) *progressive* (contribute to human progress). For example, architecture is typically defined as an occupation concerned with designing and overseeing the construction of buildings to meet the requirements of building codes, the contract with the client, and generally accepted aesthetic standards. There is no requirement that tomorrow's architecture be better than today's.

If technology consists of a set of relatively reliable procedures for managing some aspect of the world, then a flight of fancy cannot be technology. Indeed, it must be *pseudo-technology* insofar as what it uses is necessary to claim that reliance. For example, an invention cannot be useful if its usefulness would violate the third law of thermodynamics (entropy). That law is too fundamental to technology for one invention to overcome the evidence supporting it gathered over several hundred years.

Although engineering is generally as hostile to pseudoscience as science itself is, it is not always hostile in the same way. For example, engineering may treat as fact claims for the safety of a design standard when the only support those claims have is *engineering judgment* (engineering's equivalent of common sense). Some of the standards in building codes, the highway's "Green Book," and the like

have no other support. Engineering judgment is a standard of proof more rigorous than common sense but less rigorous than what science or technology typically requires. Generally, engineers appeal to engineering judgment when no one is willing to pay to do the research necessary to confirm or disconfirm the hypothesis or the necessary research would be unethical or the research is otherwise impractical, and the projects require a common standard more rigorous than common sense. Engineers do not use engineering judgment to support a theory just because they regard engineering judgment as sacrosanct or even particularly reliable. Engineering judgment is merely better evidence than nothing (common sense or a flight of fancy).

A hypothesis receives confirmation when research, done according to the appropriate discipline, uncovers evidence in its favor (i.e., facts encouraging reasonable trust in the hypothesis); it receives disconfirmation from the uncovering of disfavorable evidence. The discipline determines what counts as evidence for or against.

The distinction between *pure science* and *practical science* is hard to state briefly in a way helpful for understanding the role of theory in engineering. Most sciences seem to have both a pure side and a practical side. The pure side seeks understanding of the world for its own sake, for example, knowledge of how to construct a chemical compound with certain elements arranged in a certain way. The pure theorist is unconcerned with how that understanding will be used outside their office, laboratory, conferences, and journals. In contrast, the practical scientist seeks understanding for the sake of something *in the world*, for example, a new explosive for lunar mining. Whether any scientist is *absolutely pure* or *absolutely practical* is a question for empirical research. The same may also be true for some technologists, such as architects or computer scientists, but not for engineers. There does not seem to be a use for a distinction between *pure engineering* and *practical engineering*. Engineering (and engineers) are practical by definition, and so are engineering theories.

At least since about 1800, engineers have defined engineering as aiming at the improvement of humanity's material condition. "Humanity's material condition" connects engineering theory with physics, chemistry, and other material sciences, while *improvement* connects engineering

not only with economics, psychology, and other human sciences but also with practical reason (prudence, morality, and ethics). No activity can be engineering if it only seeks understanding for its own sake.

Engineering Theory and Technology

Though the exact meaning of the term *theory* seems to vary with the discipline studied, in general, theory is distinct from fact. So, for example, when lawyers talk about *the theory of a case*, they mean the story used to arrange the facts of a case into a reasonable or persuasive order, including *facts* still to be established. For lawyers, the facts are what they think they can prove in court. A good theory of the case is a way to present the facts likely to win a favorable judgment. Unlike theory in the sciences, the lawyer's use of *theory* in *theory of a case* has no (necessary) connection with truth. (Lawyers are rhetoricians, not scientists.)

Like most important terms, *technology* also has several meanings. In one, technology means the sum of processes, skills, and machinery used in the production of material goods (anything from books to bombs). The sum of processes, sources of power, skills, and machinery used in the production of immaterial goods, such as poetry or a social event, is not, as such, technology. So, for example, someone might properly say of the gradual shift from watches (time pieces) with mechanical innards to watches with electronic innards that "digital technology is replacing mechanical technology." However, one cannot (in this sense) properly say of the Protestant Reformation's switch from the Latin liturgy of the Catholics to a vernacular liturgy of the Protestants that "during the 1500s the technology of Protestantism replaced the technology of Catholicism in the sects of the Reformation." Being largely immaterial, the religious practices in question are not technology. It is better to speak of them as "the culture of Protestantism" and "the culture of Catholicism."

In another sense, however, *technology* refers to any system by which humans and machines work together to accomplish some practical end, whether material or not. This sense of *technology* includes culture insofar as, for example, religion relies on machines for its special foods, building designs, procedures for circumcision, ways to remember prayers, and so on.

In yet another sense, *technology* refers to systematic investigation or knowledge of technology in one of the other common senses of that word—as in such terms as *institute of technology*. Technology (in this sense) is the study of technology (techne or techniques) in one or more of the following senses of *study*: (1) passing knowledge (including knowing-that, knowing-how, knowing-when, and knowing-why) from those who have the relevant knowledge to those who do not (training or education), (2) the discovery of new knowledge of the relevant sort (research), and (3) the invention of new artifacts or new ways of producing artifacts (design).

While much technology (in one or another of these three senses) relies on scientific knowledge (natural or social), much does not. For example, the dyeing of fabric was industrialized long before the chemistry of dyeing became a science. During that prescientific age, knowledge of dyeing was largely the secret of this or that trade. That is one reason to think that technology is not just applied science.

That, of course, is an historical reason. It leaves open the question whether all technology might be applied science in the more radical sense that all technological knowledge is logically reducible to one or more of the sciences—in principle if not in practice. The answer to that question seems to depend in part on what counts as science as well as what counts as technology and reducibility. If *science* means natural science (the disciplined understanding of nature), then there is good reason to reject the claim that this or that technology relies (or must rely) entirely on one or more of the sciences. Science, in this sense, is an attempt to understand the natural world in a certain way (according to one or more scientific disciplines); but technology is, at a minimum, an attempt to change the world in a certain way. Even a true and complete understanding of nature cannot provide instructions, guidelines, or the like for changing the world—in however small a way.

There are, of course, less strict ways to understand *science* that include systematic understanding, explanation, or knowledge of nonnatural fields (artifacts or processes) as well as of natural ones. A field is nonnatural insofar as its facts depend in part on what ordinary humans do (or think) rather than on what researchers do (or think). Thus, the human sciences (history, psychology,

sociology, etc.) must be at least in part nonnatural because, for example, history is in part about what people thought—their purposes, motives, and so on. The nonnatural sciences must be *practical* insofar as they include systematic understanding of how to accomplish certain practical ends. Is this systematic understanding the link between the contemplation of science and the activity of technology?

Among the nonnatural sciences are *computer science*, *medical science*, and *political science*. Computer science is both about mathematics (or symbolic logic) and business practices (such as the processing of job applications or census data); medical science, both about how the human body reacts to various diseases and about which medications may be prescribed for which disease; and political science, both about how government works and about what politicians may do to stay in office. But even if *science* is stretched to include the study of such practical sciences (in any or all of the three senses of *study* identified earlier), technology is still not merely the application of the sciences to the world's practical affairs, such as ordering replacement parts for a stamping press or deciding that the wing on an airplane is flexible enough for ordinary use. Science, even applied or practical science, only identifies ways the world may be changed. No science recommends a way the world should be changed. The scientist who wants to order replacement parts or say that an airplane wing is safe enough for flight must go beyond the sciences and act as a technologist.

Engineers and the Philosophy of Technology

For our purposes, the philosophy of technology is the attempt to understand the reasonableness of technology. The reasonableness of a fact, activity, or practice depends (in part at least) on its definition. Definition can be difficult. For example, the philosophy of technology has long debated which *values*, if any, *technology* includes or should include in its definition.

The traditional view seems to have been that technology is a *pure means* for rational agents (persons who are prudent but not necessarily moral). Thus, the slogan “Guns don’t kill people; people kill people” treats guns as a pure means.

The source of the bad effect of guns is the people who use them—both rational agents and those suffering from one or another mental disability. If guns were not available, people would find other means to kill—a kitchen knife, a cast iron frying pan, a rock. Guns have no agency of their own.

A more recent view is that *intrinsic values* are (or can be) designed into a technology whether the designers know it or not. This treats artifacts, such as guns, as having an agency of their own. On this view, the proper response to the slogan “Guns don’t kill people; people kill people” is “People seldom kill people on their own; more often, guns make the killing easier. Take away the guns and there will be less killing.” The guns in question are in the causal chain just as much as people are (when a gun is used).

Another recent view rejects *rational agents* as the typical participants in technology, substituting *the reasonable person*. Reasonable persons are, of course, prudent (most of the time at least) and have a basic knowledge of the world. They can plan. Their plans typically suit their desires. And their acts typically suit their plans. In these respects, reasonable persons are much like rational agents. The most important difference is that reasonable persons typically care about others while rational agents do not (especially the rational agents of economics). Another important difference is that reasonable persons recognize morality as a desirable constraint on everyone’s conduct, not as a mere cost of doing business, as rational agents do. These two differences are empirical (as well as theoretical). There is considerable evidence that most people do in fact care about others, not only family and friends but also people more distant, such as victims of oppression in other countries, though distance does seem to reduce the strength of their caring. Most people probably care most about themselves and their immediate family, but people who care only about themselves (the merely prudent) seem to suffer from a character flaw. Among the reasons for people to care about morality is that they care about others, including children and the disabled, and morality protects such people.

Engineers seem to design for reasonable persons who are not always at their most reasonable. They try to foresee distraction, exhaustion, ignorance, and the like and seek to prevent such disabling factors from causing *operator error*. Indeed, engineers

typically try to make the artifacts or processes they design *foolproof*—or at least as foolproof as budget, organizational policy, regulation, and the like allow. Other technologists seem to work to a different standard, for example, protecting the reasonable person (but not the fool) from suffering or causing serious harm. Moral theory, rational decision theory, and human factors analysis have an important place in setting design criteria and in evaluating them after the design has entered the world as an artifact or process.

There is, at any time, a practical upper limit even to engineering's *safety factors*. Typically, that limit is set by the market. If unit cost of an artifact or process (after including the cost of design, production, operation, maintenance, disposal, and reasonable return on investment) exceeds what people are willing to pay (*demand*), there will generally be no market for that artifact or process. The exception will be for those artifacts (or processes) that have subsidy lowering the market price enough to be attractive. The most common subsidies come from taxation (to support parks, roads, public schools, and other government services) and grants from private charities (supporting libraries, museums, hospitals, and other nongovernmental agencies). Such exceptions from market control are typically defended as correcting for *market failure* (e.g., for the fact that everyone benefits more from the practice of sharing roads built at public expense than they would from the practice of letting each landowner build his own stretch of road and take tolls from users).

Most new artifacts (or processes) that engineers design replace similar artifacts. The new artifact should be better than the old (i.e., improve humanity's material condition). Among the factors making an artifact better than what it replaces is lower cost (because of use of cheaper but equally reliable materials, use of less material, or a better manufacturing process), other improvements in *technology* (such as more functions, lower weight, and easier disposal), and better support (design making repair or recycling easier). If the cost of an improvement in design is, all else equal, negligible, an engineer should automatically include the improvement in the design to raise the *state of the art*. If the cost of the improvement is significant, engineers should propose it to the organization

for which they work. When, however, substantial improvement competes with substantial cost, engineers can rarely decide the question by appeal to economic theory or engineering judgment. Determining whether the improvement should be made seems to require a social decision—or at least a decision involving more than engineers consider (or should consider) to be within their competence. The question of when a decision is best left to an expert and when it is so political that the public should have a part has been a lively subject in political theory since the ancient Greek philosopher Socrates.

When engineers speak of *a design*, they typically mean the documents giving instructions for carrying out some major undertaking, say, the building of a new cracking tower or manufacture of a new microchip. This is design with a *big D*. Less often, engineers use *design* to mean the instructions for carrying out smaller projects, such as a proposal for replacing lead pipe with copper in the water delivery system of a 12-floor apartment building. Most designs, even small-d designs, typically rely on engineering theory of one sort or another.

Theory in Engineering

Though the term *theory* has many senses, those senses generally have this much in common: there is a *story* (a set of hypotheses) on which to hang facts, especially new facts that experiment, observation, or reasoning reveals or might reveal. The understanding that a scientific theory promises consists not primarily in the new facts it leads to but in the proposed *story* that the facts, old and new, make increasingly plausible, increasingly coherent, and increasingly suggestive of new hypotheses. Much the same is true of engineering theory—except that its research typically identifies directives that ought to be followed (or not), for example, the *fact* that *Y* is best for some purpose.

Theories may, then, be thought of as guides to research (suggestions possibly leading to interesting hypotheses or whole research agendas or as organizers of (supposed) facts (facts that support one or more hypotheses). Most theories belong in both categories. Calling a certain set of hypotheses *a theory* usually signals a research agenda, one that may, when followed, confirm or disconfirm

the theory. But once a theory has been confirmed enough to become central to an important science or technology, it *dies and goes to heaven*, that is, it ceases to generate interesting research concerned with its confirmation or disconfirmation and instead contributes to science and technology in less direct ways, especially by offering a useful perspective on a family of research agendas.

Insofar as that is true, *engineering theory* has a weaker analogy with theater than *scientific theory* does. In a theater, the audience (typically) observes the actors. It does not interfere with them. The actors are *nature*, that is, that part of the world that the science in question crafts its theories to understand. In engineering, however, theory is (in part at least) about how to interfere with the actors (typically the human beings and natural processes affecting human artifacts or processes) to achieve some material improvement in human life. For example, engineering can reduce the risk of accident by designing an appropriate schedule for inspecting an airplane's wing for signs of fracture. Designing the appropriate schedule will, of course, require engineers to consider at least two factors foreign to natural science: cost and safety. An appropriate schedule will not waste resources by, say, starting inspections of the wing too soon or by using uninformative tests. An engineer's schedule will, however, typically include a *safety factor*, that is, enough redundancy in the required inspections to allow for the effect of unusual temperatures on test devices, the occasional exhaustion of inspectors, and so on.

Because controlling cost and assuring a large enough safety factor are competing considerations, even designing a simple inspection schedule for an airplane's wing may require a complex theory, one directing how to determine and weigh moral considerations such as waste and safety, as well as facts about the nonhuman world, such as metal fatigue. Neither the natural nor social sciences will have much to contribute to the relevant theory. It is engineering.

Michael Davis

See also Fact Versus Theory; Hypothesis Testing; Knowledge; Laws (Scientific); Rationality; Received View of Theories; Speculation; Understanding; Values in Science

Further Readings

- Aubin, D. (2017). Ballistics, fluid mechanics, and air resistance at Gâvre, 1829–1915: Doctrine, virtues, and the scientific method in a military context. *Archive for History of Exact Sciences*, 71, 509–542. <https://doi.org/10.1007/s00407-017-0195-6>
- Brey, P. (2010). Philosophy of technology after the empirical turn. *Techné: Research in Philosophy and Technology (Winter)*, 14, 36–48.
- Chebotareva, E. E. (2019). Engineers: Bridging the gap between mechanisms and values. *Social Epistemology*, <https://doi.org/10.1080/02691728.2019.1695008>
- Curry, O. S., Chesters, M. J., & Van Lissa, C. J. (2019). Mapping morality with a compass: Testing the theory of 'morality-as-cooperation' with a new questionnaire. *Journal of Research in Personality*, 78(February), 106–124.
- Gordon, J. E. (2020). *Structures, or why things don't fall down*. Hachette Books.
- Kanemitsu, H. (2018). New trends in engineering ethics—A Japanese perspective. In A. Fritzsche & S. J. Oks (Eds.), *The future of engineering, philosophy of engineering and technology*. Springer.
- Koen, B. V. (1985). *Definition of the engineering method*. American Society of Engineering Education.
- Layton, E. (1971). *The revolt of the engineers*. Case Western Reserve University Press.
- Mitcham, C. (1994). *Thinking through technology: The path between engineering and philosophy*. The University of Chicago Press.
- National Academy of Sciences, Institute of Medicine (2001). *Science, evolution, and creationism*. National Academies Press.
- Noble, D. E. (1977). *America by design*. Alfred A. Knopf.
- Pahl, G., & Beitz, W. (1988). *Engineering design: A systematic approach* (Arnold Pomerans & K. Wallace, Trans.; K. Wallace, Ed.). The Design Council London.
- Petroski, H. (1996). *Invention by design: How engineers get from thought to thing*. Harvard University Press.
- Suppe, F. (2000). Understanding scientific theories: An assessment of developments, 1969–1998. *Philosophy of Science*, 67(September), S102–S115. <https://doi.org/10.1086/392812>
- Vincenti, W. G. (1990). *What engineers know and how they know it*. Johns Hopkins University Press.
- Zima, P. V. (2007). What is theory? Cultural theory as discourse and dialogue. Continuum. [Translated from: *Was ist Theorie? Theoriebegriff und Dialogische Theorie in der Kultur- und Sozialwissenschaften*. A. Franke Verlag, 2004].

ENGINEERING, HISTORY OF

The history of engineering represents essential knowledge for today's engineers, showing how engineering ideas, training, and practices evolved in different countries and cultural contexts. Multiple intellectual, social, and political factors shaped centuries of engineering, including the formalization of education, professionalization of expertise, corporate and government connections, wartime demands, and pressures of economic competition. Recent decades have highlighted ongoing issues, including engineering and ethical responsibility, modernization and training accreditation, and strategies for diversifying engineering perspectives by widening access to women and others historically underrepresented.

Early History of Engineering

Engineers existed even before that term came into use, practitioners who developed practical tools, buildings, and infrastructure. Across early Africa, the Mediterranean, Asia, and elsewhere, innovators (names lost to time) created and improved agricultural implements, food storage systems, early textile-making and metalworking, new boat forms, and navigation techniques. Creators planned and managed construction of Stonehenge and other prehistoric monuments, Egypt's pyramids, Greek and Mayan temples, and defensive walls in cities across cultures. Skilled leaders supervised the movement of thousands of tons of material (without modern mechanization) to create massive stonework, often with precise angles or aligned with the stars. Planners designed and managed construction and maintenance of Inca road networks, early Middle Eastern and Chinese canals for irrigation and transport, and Roman systems of roads, sewers, and aqueducts, all stretching across hundreds of miles.

Similarly in medieval Europe, master masons and other men, with no formal engineering degrees but years of hands-on experience, directed design and erection of soaring Gothic cathedrals. Well-designed waterwheels powered manufacture of cloth, flour, iron, and other goods, driving increased commerce and economic growth. In China, specialists refined and expanded production of iron,

steel, and bronze, metallurgical knowledge applied to make more efficient plows and other farm tools, stronger weapons, and ceremonial objects such as bells. Other innovators in China created faster ships and magnetic compasses for long-distance exploration and trade, plus early kites and gliders, silk-manufacturing tools, and fireworks, including gunpowder rockets and projectile weapons. In the Islamic world, Ismail al-Jazari created and described numerous the so-called *ingenious devices*, including clocks and automata for amusement.

European Military Engineering

The word *engineer* originated around the 1300s, commonly referring to men who specialized in devising military constructions and derived from Latin words meaning "to create or invent." Warfare in early modern Europe offered engineers employment in designing new weapons and fortifications for wealthy patrons. Leonardo da Vinci (1452–1519) spent years inventing improved catapults, cannons, and muskets for Italian and French rulers, while also designing canals, bridges, industrial machines, cathedrals, and fortifications. Leonardo's impractical but bold ideas for armored tanks, attack submarines, and flying machines made his name synonymous with brilliant creativity. But Leonardo built on the work of earlier engineers, including Germany's Konrad Kyeser (1366–c. 1405) and Italy's Mariano di Jacopo, known as *Taccola* (1382–c. 1453), who wrote books conveying specialized knowledge of the era's arms-making and other machinery. Italy's Francesco di Giorgio Martini (1439–1501) revolutionized the design of military fortifications for the age of artillery. Like Leonardo, he combined work in engineering and art; both men applied metalworking expertise to make military cannons and also decorative statues. Italy's Vannoccio Biringuccio (1480–1539) and Germany's Georgius Agricola (1494–1555) published masterworks on mining engineering and metallurgy. Agostino Ramelli (1531–c. 1610) served as a military engineering consultant in both Italy and France and published many machine designs, especially hydraulic devices for raising water. Dutch engineers such as Cornelius Vermuyden (1595–1677) specialized in swamp drainage and water management for land reclamation.

Europe's link between engineering and state military priorities continued through the 1600s, highlighted by the fortification and siege machine designs of France's Sebastien de Vauban (1633–1707). By the 1770s, French leaders had established the École Nationale des Ponts et Chaussées, an institution to foster new generations of experts in road and bridge engineering. Educators wrote the textbooks for their new curriculum, focused on mathematical theories of structures and scientific understanding of materials. In the late 1700s and early 1800s, France opened several new institutions to teach engineering, machinery, and mining, including the famous École Polytechnique of 1792. Graduates primarily served in France's military, linking technical command to state power, but others took positions in French industry and urban planning. French engineering theory fostered key developments in thermodynamics, hydrodynamics, vibration, and statics. Due to the era's links between technical knowledge and military needs, the military academy at West Point became the first school in the United States to teach engineering, feeding into the U.S. Army Corps of Engineers. Graduates also led in creating roadways, bridges, canals, and harbors for a rapidly growing nation.

Industrial Revolution Engineering

Well into the 1800s, many men continued to fulfill engineering duties without academic qualifications. Britain's Industrial Revolution relied on experts trained through apprenticeship, focusing on practical knowledge. British textile mills, other factories, and iron-making blast furnaces needed dependable power sources. Thomas Newcomen (1664–1729) transformed earlier steam devices to create an atmospheric engine useful for pumping water out of deep mines. Instrument maker James Watt (1736–1819) then improved Newcomen's design by adding double-acting mechanisms and a separate condenser. His more efficient and versatile steam engine promoted expansion of large-scale manufacturing, raised productivity, and led to faster modernized road, rail, and boat transportation. Among other Industrial Revolution engineers, John Smeaton (1724–1792) made key mechanical engineering improvements to both waterwheels and steam engines. Blacksmith

worker Henry Maudslay (1771–1831) led in machine-tool innovation, creating precision measuring devices, inventing metal-cutting lathes to produce standardized screws, and starting a firm that manufactured steam engines and gave young mechanical engineers on-the-job training.

The Industrial Revolution's growth of trade also rested on advances in transportation engineering and infrastructure construction. Scottish engineer Thomas Telford (1757–1834) rose from a stonemasonry apprenticeship to become the first president of Britain's Institution of Civil Engineers. Telford won renown for designing important British canals, improved roads, and bridges, including his 1826 structure crossing the Menai Straits, then the world's largest suspension bridge. Marc Isambard Brunel (1769–1849) developed tunneling technology to construct the Thames Tunnel, the first passage under a major river. His son Isambard Kingdom Brunel (1806–1859) served as chief engineer for some of England's earliest significant railroads, including the Great Western Railway, surveying routes, selecting rail gauge and rolling stock, and designing tunnels, bridges, and stations. Brunel's civil engineer expertise extended to marine engineering; he created the era's largest passenger and cargo steamships, including the SS *Great Eastern*, used to lay the first successful transatlantic telegraph cable in 1865–1866.

Engineering Systems, Education, and Professionalization in the Late 1800s

The 19th century brought continued expansion of national and global transportation routes, connected with ongoing growth and mechanization of manufacturing. Engineers played central roles in facilitating those developments. Alexander Graham Bell (1847–1922), Thomas Edison (1853–1937), Nikola Tesla (1856–1943), Guglielmo Marconi (1874–1937), and numerous others developed networks for electrification, telephone communication, and radio broadcasting. Realizing ambitious technical goals required the creation of complex systems. While many experimenters of the 1800s had investigated electricity and built arc-lighting systems, Edison had to do more than just invent a more practical incandescent bulb. Working with a team of engineers, scientists, and

mechanics at his Menlo Park, New Jersey laboratory, Edison created better electric lamps, dynamos, wires, and conduits. Edison then established companies to manufacture all those components vital to making electrification feasible and in 1882, opened Pearl Street Station in New York City, the first permanent commercial system of centrally generated electric light and power. Edison then set up hundreds of other direct-current systems across the United States and beyond, which hired hundreds of young technicians and engineers. Tesla worked for Edison before starting his own career that challenged Edison by building alternating-current electric technology. Before entering automobile making, Henry Ford gained valuable experience working for Detroit's Edison Illuminating Company.

The clear potential for new large-scale systems and machinery to have large-scale economic, political, and social impact encouraged the formalization of engineering education. Political, business, and academic leaders hailed expansion of formal training as essential in opening possibilities for economic growth and technical progress. In the United States, a key development came with Congress's passage of the 1862 Morrill Land-Grant Act to finance establishment of public colleges focused on useful training in the mechanic arts and agriculture. By offering accessible engineering education, states hoped to support entrepreneurs opening new businesses and nurture a population of male middle-class breadwinners. By the late 1800s, land-grant schools across the country combined with private schools, such as the Massachusetts Institute of Technology, to graduate hundreds of engineers each year. While many engineering curricula still stressed hands-on machine-shop work and other applied practice, the formalization of training set new standards.

Germany's Technische Hochschulen, established starting in the 1800s, combined influence of French engineering training with practical emphasis. These institutions and their graduates earned high international regard, especially in connection with Germany's world famous chemical and dye industries. Other countries similarly supported growth of formal engineering education as a tool for advancing state economic and political interests. British engineering training took root in the late 1800s and early 1900s at "red brick

universities," not coincidentally located in manufacturing centers such as Liverpool and Leeds. Spain in the 1800s established technical schools for mining, forestry, agriculture, and other fields considered essential to the civil service. Rigorous examinations and long hours of study reinforced the institutional impression of a meritocratic ideal. Russia's czars supported engineering education as a pathway to modernization. St. Petersburg, Kiev, Warsaw, and other major cities across the Russian empire set up new technical training schools in this era, modeled on polytechnic institutions in Britain, France, and Germany.

Across Europe, the United States, and elsewhere through the 1800s and into the 20th century, many men (and almost exclusively men) without college degrees continued to enter and succeed in engineering work. Those workers picked up mastery through more informal training and hands-on experience. But by the early 1900s, a growing number of large companies in particular sought to employ engineers with academic qualification. Ambitious young men hoped that engineering education could provide a secure career and upward social mobility.

Not coincidentally, a wave of professional organization in this same period allowed engineers to define themselves by discipline. In the United States, the American Society of Civil Engineers was formed around 1870, soon followed by groups for mining engineers, mechanical engineers, electrical engineers, and chemical engineers. These new societies set criteria for membership, which in turn codified expectations of experience. While politicized issues of ethics and standards sometimes raised controversial dissension, society conferences, publications, and decisions gave each discipline a voice to assert its interests with both educators and employers. Engineering groups also helped set safety standards, building codes, and quality monitoring. Nations increasingly passed laws governing the accreditation, testing, licensing, and registration of professional engineers.

Engineering Through World War II

With the expansion of industry and establishment of hierarchical large enterprises, engineers found new roles as corporate servants. Many shifted careers into management, with some reaching the

upper ranks of major firms. The early 1900s brought the creation of industrial research centers such as Bell Telephone Laboratories, Westinghouse, General Electric, DuPont, General Motors, and many others. Those institutions financed engineering and scientific research, meant to translate into a production and marketing edge. Company leaders expected tangible rewards in efficiencies and innovation in new products. Meanwhile, engineers based inside big businesses had to balance professional independence against expected loyalty to an employer, sometimes facing tough decisions about whistleblowing and concern for the public good.

The growth of big business also promoted the growth of new engineering specialties. In 1911, mechanical engineer Frederick Winslow Taylor (1856–1915) published his principles of *scientific management*, strategies for rationalizing factory design and supervising workers to maximize productivity. Although some labor elements resented and resisted the manipulation and threats of deskilling, business leaders embraced the emerging discipline of industrial engineering as promising profitable efficiencies. Henry Ford's success in institutionalizing the automobile assembly-line drew worldwide attention after 1913, and engineers built on the system of *Fordism* with further innovations in management of mass production.

In the early 1900s, as in earlier centuries, national leaders used powerful military, communications, and building technologies to facilitate conquest, colonization, and control. Numerous European and U.S. engineers worked with politicians and administrators to supervise the construction of roads, railways, telegraph lines, and water systems in remote areas of Asia, Africa, and South America. American engineers assisted with military missions in the Philippines in the early 1900s. Their careers embodied rhetoric that linked new technologies to White political dominance, and the supposedly natural advance of *civilization*. Promises of material wealth and improved well-being allowed colonizers to rationalize the harsh suppression of native peoples and serve Western political and economic interests.

Alongside the early 20th-century start of the automobile age, Orville and Wilbur Wright's 1903 success with powered flight at Kitty Hawk, North Carolina, opened up the airplane age and

ultimately the field of aeronautical engineering. Civil engineering drew widespread public attention; impressive 19th-century projects, such as the Brooklyn Bridge, completed in 1883, and 1889's Eiffel Tower set the stage for even more impressive construction, including the Panama Canal (1914), Empire State Building (1931), the Hoover Dam (1936), and the Golden Gate Bridge (1937). In the 1930s, Germany began creating its high-speed autobahn network, while engineers in the Soviet Union created the massive steel production center of Magnitogorsk. Nations invested in engineering expertise; for example, the interwar U.S. government set up the National Bureau of Standards and the National Advisory Committee for Aeronautics (precursor of the National Aeronautics and Space Administration) to foster engineering and scientific research and development.

Twentieth-century conflicts provided employment for engineers on all sides to work with national military agencies and weapons manufacturing companies in devising new weaponry, defensive technologies, and other wartime applications. World War I's destructiveness was tied to the development of powerful naval dreadnoughts and submarines, chemical weapons, tanks, and military aviation. World War II engineers and scientists created radar, rocketry, and the atomic bomb, all feeding into the technological drives of the Cold War arms race and space race between the United States and the Soviet Union. World War II also accelerated computer development in the United States and Britain, setting up postwar development of new programming languages, transistors, and microprocessors. Computer engineering thrived as marketers found new public and private customers for large-scale data handling and automation. The postwar decades also supported growth in the fields of nuclear engineering and aerospace engineering in the United States, Soviet Union, and Europe.

Late 20th- and Early 21st-Century Developments and Issues in Engineering

The post-World War II culture of technical education shifted toward the so-called engineering science, deepening emphasis on advanced mathematics, theoretical analysis, and scientific techniques. Graduate work at elite schools drew

top students from around the globe, eager to obtain the imprimatur of the most respected degrees. Further intellectual evolution fostered the rise of cross-disciplinary studies and entirely new fields, including expansion of medical engineering and bioengineering. Practices of everyday engineering work modernized, shifting away from slide rules to calculators, away from mechanical drawing to computer-aided design. But by the late 20th century, concern that education had become overly abstract led many engineering classes to move toward reinstating hands-on practice and real-life applied problem-solving. Employers expected engineering hires to master not just technical knowledge but also skills of effective communication, global teamwork, and leadership.

The 1960s and 1970s connected engineering with both highs and lows, as admiring crowds worldwide celebrated the first moon landing, but the Vietnam War and other conflicts showed the ongoing human costs of increasingly destructive weaponry. The growing environmental movement called attention to the costs of pollution and over-exploitation of natural resources. The 1980s and 1990s underlined perceptions of engineering as politically and economically crucial to national competitiveness, connecting technological expertise to productivity of manufacturing, services, and communication.

Meanwhile, the late 20th century and early 21st century brought new landmarks and new issues in engineering. Completion of impressively large global structures and infrastructure continued, including China's Three Gorges Dam (1994), the British-French Channel Tunnel (1994), and ever-taller skyscrapers. The supersonic airliner Concorde, the space shuttle, Mars rovers, the International Space Station, and other high-profile engineering and scientific landmarks drew public attention. Generations of consumers took for granted immediate access to information, messaging, and entertainment made possible by the latest smartphones, Wi-Fi access, even smart watches and virtual reality. High-tech development thrived, as the mindset of California's Silicon Valley spurred attempts to seed similar *innovation hubs* across the United States, Europe, and elsewhere. Nations including Japan, China, India, and more recently in the Middle East cultivated rapid expansion of engineering training as a shortcut to

desired economic and political transformation. Computer-driven shifts in production, communication, distribution, and sales spread associations of modernization with abundance and choice in food, clothing, housing, and personal devices. Popular culture made engineers, entrepreneurs, and innovators into new celebrities; many admired such figures as Bill Gates, Steve Jobs, Elon Musk, and Mark Zuckerberg, even as critics denounced Silicon Valley's mindset of "move fast and break things."

Innovations in technology came alongside questions and concerns about costs, benefits, and risks. Lack of investment in disaster prevention and infrastructure maintenance led to unnecessary deaths in incidents every year. Technological changes across workplaces roused fears of displacement and deskilling among both well-educated and manual employees. Easy access to online information also opened the gates to misinformation and fake news, plus difficult international debates over privacy, piracy, bias embedded in algorithms, and potential abuses of power. Engineering continued to be both a tool and a pawn of nations and regions jockeying for political and economic dominance. Secrecy, spying, and competition complicated work in nuclear, manufacturing, and computer engineering. The military-industrial complex's interest in cutting-edge technology continued to provide employment for tens of thousands of engineers worldwide, even as the power of weaponized drones and robotic warfare raised new moral hazards.

Economic and political inequities within and across nations both reflected and reinforced technological gaps; for instance, advantaged patients in well-off countries could benefit from robotic surgery, advanced diagnostic tests, and other sophisticated medical tools, while disadvantaged citizens still suffered from contaminated water, food shortages, and preventable diseases. Public and professional attention to such problems spurred debate about the social responsibility of engineering. The 1970s' *appropriate technology* movement analyzed environmental ethics and cultural fit. In later decades, advocacy groups such as Engineers Without Borders and Engineers for a Sustainable World aimed to combine professional expertise with an outward-looking activism. Concerns about climate change drove growing

21st-century investment in green design and renewable energy engineering, alongside debates about carbon sequestration technology and future radical geo-engineering. Yet as for centuries, the pathway of technological development was never predetermined. Human choices, uncertainties, cultural constraints, and political disagreements shaped the terms of engineering work and complicated decision-making about desired outcomes. Individual and social conditions even affected which matters were defined as engineering problems to be solved and how.

Meanwhile, 20th- and 21st-century campaigns aimed to open more opportunities in engineering for women and other historically underrepresented groups. For centuries, technical careers had remained almost universally the province of White men. Engineering's historic ties to the military, field projects, and corporate ranks blocked access for most women. Rhetoric promoted engineering as a masculinized culture, valorizing daring heroes who built railroads across deserts, roads through jungles. Building and model construction toy kits explicitly encouraged boys to enjoy tinkering, promising that hands-on play would lead young men into engineering jobs.

During the early 1900s in the United States, Canada, and Europe, a handful of women had pressed ahead to earn engineering degrees at the few schools willing to accept them. Many faced isolation, skepticism, ridicule, or opposition from parents, high-school teachers and counselors, and both male and female classmates. Engineering faculty, deans, and employers often openly doubted whether female students could succeed. For decades, major engineering organizations such as the American Society of Civil Engineers refused to accept women. Lillian Gilbreth, who had worked hard to establish a strong reputation in industrial engineering, still confronted an uphill battle to win full membership in the American Society of Mechanical Engineers.

In the 1950s, women still comprised less than one percent of engineers in the United States and many other countries, though substantially more in the Soviet Union. In Western nations, the small but growing ranks of female engineers organized to support each other and encourage more young women to consider engineering careers. A handful

of female engineering school faculty and staff, together with male allies, offered psychological and practical support in what remained a frequently chilly and even toxic climate. Female students who persisted to earn degrees confronted extra hurdles in finding employment. Hiring committees questioned women about when they would get married, doubted whether a female engineer could think about mechanical details "like a man," and offered salaries distinctly lower than men's.

In the late 1960s and 1970s, feminist movement mobilization led women in engineering studies and careers to mobilize to confront those hiring biases, sexual harassment, and the challenges of juggling work and family. By the 21st century, major engineering programs, professional organizations, and other leaders in the United States, Europe, Asia, and elsewhere mounted and joined campaigns for diversifying engineering. Global corporations, governments, schools, and other parties launched multimillion-dollar initiatives to expand access to engineering and computer education for more youngsters from previously underrepresented groups. Female engineers continued to report harassment and discrimination from employers and coworkers in Silicon Valley and other job sites. For complex reasons, women's representation in engineering still varied dramatically between different countries and across technical disciplines. But discussions of social justice and embedded bias aimed to reshape engineering culture for a modern age.

Amy Sue Bix

Further Readings

- Adas, M. (2006). *Dominance by design: Technological imperatives and America's civilizing mission*. Harvard University Press.
- Billington, D. P. (1996). *The innovators: The engineering pioneers who made America modern*. Wiley.
- Bix, A. (2013). *Girls coming to Tech!: A history of American engineering education for women*. MIT Press.
- Brown, J. (2000). Design plans, working drawings, national styles: Engineering practice in Great Britain and the United States, 1775–1945. *Technology and Culture*, 41(2), 195–238.
- Calhoun, D. (1960). *The American civil engineer: Origins and conflict*. Harvard University Press.

- Calvert, M. (1967). *The mechanical engineer in America, 1830–1910: Professional cultures in conflict*. Johns Hopkins University Press.
- Hill, D. (1996). *A history of engineering in classical and medieval times*. Routledge.
- Kranakis, E. (1997). *Constructing a bridge: An exploration of engineering culture, design, and research in nineteenth-century France and America*. MIT Press.
- Layton, E. (1986). *Revolt of the engineers: Social responsibility and the American engineering profession*. Johns Hopkins University Press.
- Marcus, A. (Ed.). (2005). *Engineering in a land-grant context: The past, present, and future of an idea*. Purdue University Press.
- Nienkamp, P. (2015). Engineering national character: Early land-grant college science and the quest for an American identity. In A. Marcus (Ed.), *Science as service: Establishing and reformulating American land-grant universities, 1865–1930*. University of Alabama Press.
- Oldenziel, R. (1999). *Making technology masculine: Men, women, and modern machines in America, 1870–1945*. Amsterdam University Press.
- Rae, J., & Volti, R. (2001). *The engineer in history* (Rev. Ed.). Peter Lang.
- Reich, L. (1985). *The making of American industrial research: Science and business at GE and Bell, 1876–1926*. Cambridge University Press.
- Reynolds, T. S. (Ed.). (1991). *The engineer in America: A historical anthology from technology and culture*. University of Chicago Press.
- Reynolds, T. S. (1992). The education of engineers in America before the Morrill Act of 1862. *History of Education Quarterly*, 32(4), 459–482.
- Seely, B. (1993). Research, engineering, and science in American engineering colleges, 1900–1960. *Technology and Culture*, 34(2), 344–386.
- Wisnioski, M. (2012). *Engineers for change: Competing visions of technology in 1960s America*. MIT Press.

ENVIRONMENTAL STUDIES

Environmental Studies as an academic discipline is relatively new, with higher level degree options only available for the last few decades. Ironically, for most of human intellectual history, academics did not fragment into specialization within just one single discipline. With rapid increase in knowledge across the academic spectrum produced since the

Enlightenment, the trend toward specialization continues at an accelerated pace. Thus, the efforts of those who synthesize, integrate, and draw connections between disciplines now seem new. Let it suffice to say that the term *environmental studies* has only recently been coined, whereas the ideas it addresses have deep traditions. The core of environmental studies seeks to emphasize the interrelations both between natural systems and between human beings and natural systems. Unlike other academic disciplines, inherent within environmental studies is the strong ethical belief that human–nature interactions ought to be as benign as possible for both humans and nature. The fact that this is often not the case drives a sense of urgency for change in our current ways of interacting with nature that underlies contemporary environmental studies.

Because environmental studies emphasize dynamic interaction between human and natural systems, it is necessary to integrate and span across both the natural sciences and the social sciences and to include methodologies that range from extremely quantitative to highly qualitative. This wide umbrella can be perceived as both strength and weakness for environmental studies, since the breadth of knowledge acquisition must at some point sacrifice a level of depth in any one discipline. That said, one thing is certain; today's generations face complex problems at multiple scales, ranging from global climate change to local sources of pollution. But any solutions to these problems require understanding political, economic, historical, and cultural forces as well as impacts to biogeochemical systems, and often, these problems persist specifically because people with different aspects of understanding and knowledge do not talk to each other or use the same vocabulary, nor do they understand how human actions impact the natural environment.

Given the interdisciplinary and integrative nature of environmental studies as a discipline, it is common practice for environmental specialists to choose from a wide array of theories developed by other disciplines and utilize them in new or innovative circumstances toward an environmental end. Being theoretically promiscuous often results not in scandal, alienation, or outrage but in innovation, curiosity, and creative problem-solving. Theoreticians within most disciplines spend their

intellectual energies seeking to define, refine, and expand the core ideas that make a specific academic pursuit, but the discipline of environmental studies is uniquely situated to adapt theories across a variety of disciplines, expanding not at the core but at the leading edges between disciplines. In this unmapped territory that most traditionalists of any single discipline consider to be populated by heretics and renegades, entirely new and creative thinking, which transcends the constraints of any one discipline, becomes possible. Indeed, this creative and transcendent thinking is not just possible but desperately needed if we are to solve the pressing environmental issues that surround us. This entry focuses first on some background ideologies that set various foci for environmental studies and then traces some cultural developments that have resulted in the discipline we now have, several strands of which are discussed in the last five sections.

Anthropocentrism

The concept of *anthropocentrism* gives a name to the deeply embedded belief that human welfare matters most, and whatever interactions occur between humans and the natural environment become relevant only as they benefit humans. This often takes the form of material benefits that increase economic prosperity; raise the human standard of living; entail the capture, storage, and use of natural resources; and most specifically focus on improving the *condition of man* as a desired outcome. Although it may seem that this describes most people in their daily activities, it is worthwhile to note how strong a role cultural identity plays in determining whether, or to what degree, a person, group, or nation stresses anthropocentrism.

In America, anthropocentrism often coincides with other beliefs that are not necessarily the same but have overlapping agendas, such as individualism, private property, the Judeo-Christian tradition, and capitalist economic values. In order to prioritize the welfare of human beings as the basis of all activities, this usually means accruing wealth and maximizing the right of the individual to behave in any manner that improves individual conditions—regardless of impacts upon others or the natural environment. The motivation to be anthropocentric can often tie deeply with religious

beliefs predicated on being chosen as a favorite creature of God in his kingdom and/or the belief that humans were put on Earth to tame nature. Other cultures and countries have far different but equally deeply rooted views on how humans ought to interact with nature. Cultural perception of nature diverges most sharply from anthropocentrism in regions dominated by animist and tribal religions and in the presence of the Asian religions of Buddhism and Shinto.

Ethnocentrism

Similar to anthropocentric views but on a smaller scale, *ethnocentrism* prioritizes those things that improve a subgroup or ethnicity of people. In this case, humans in general matter less than one's own tribe, or group of people, but the idea is essentially the same. Efforts or actions taken to improve the standard of living, wealth, or well-being of one group would downplay or ignore residual impacts to other groups of humans or to the natural environment. This more localized prioritization of some people's benefit (usually at the expense of others) has been an underlying cause of warfare, famine, and environmental disasters since time immemorial. For instance, looking at water management, we see many regional tensions due to ethnocentric values. Israel declared a buffer zone of military occupation in southern Jordan, but it also drilled several deep wells to procure much coveted water in a desert environment—specifically at the expense of those who live in Jordan. Similarly, U.S. agricultural activities degrade the water quality of the Colorado River as it flows south into Mexico, but many Americans do not consider the burden this puts on Mexico, nor does the U.S. government seriously address appeals by Mexico to rectify the situation.

Ecocentrism

In contrast to both anthropocentrism and ethnocentrism, *ecocentrism*, also known as *biocentrism*, prioritizes natural systems and sees human beings as just one species among many. Thus, it is the ecosystem as a whole that must be considered first, when decisions are made and actions are taken. In the ecocentric view of the world, protection of natural systems is paramount, even if some

humans suffer setbacks or are constrained from maximizing their potential wealth. For instance, given the choices between clear-cutting a forest to maximize the profits gained in selling the lumber or some other land use, an ecocentric decision might be to only selectively cut trees in order to thin out the forest and sustain its long-term health. Clearly, a person cannot simultaneously be anthropocentric and ecocentric, since the values are often mutually exclusive. One cannot save a forest and simultaneously clear-cut it to develop the land. However, one dilemma that environmental studies centers on is the tendency of humans to think one thing and do another. In reality, most people live with some contradictory impulses toward both anthropocentrism and ecocentrism, and finding a balance between these conflicting impulses presents a deep challenge to daily living and short-versus long-term metrics of success, value, costs, and benefits.

Capitalism and Exploitation

The idea of commerce and trade for profit goes back thousands of years. But not until the convergent evolution of capitalism, paper currency, and changes in political theories of value and property, did we see massive transformation in the political economy and a concomitant rise in use of resources that now exists in modern capitalism. Feudalism as a system of land tenure, social organization, and economy dominated for more than 1,000 years before its dismantling, which began in Europe in the 16th century. During this time, both the numbers of people and their combined ecological effects remained stable with a relatively low impact on the environment. As agrarian capitalism gave way to mercantilism, and then industrial capitalism, theories of philosophers John Locke and Adam Smith established the fundamental ideas of industrial capitalism, and it is this system that has dominated political and economic thinking for the last 200 frenetic years of expansion.

A core idea of capitalism assumes that all participants are rational actors, who will seek to utilize the resources at their disposal to maximize their personal economic gain, and it is that instinct which drives the laws of supply and demand. It is upon these deeply embedded ideas that industrial

capitalism feeds the desire to continually accumulate more wealth. Of course, natural resources form the basic raw ingredient in this philosophy of political economy. Nowhere among the basic premises of industrial capitalism is there to be found any discussion of limits on accumulation of wealth, the consumption of finite resources, or the environmental damage done as a residue of accumulating wealth and exploiting resources. But an economic system predicated on endless exploitation of natural resources generates incentives to maximize the exploitation of such resources. And after two centuries of resource exploitation feeding the system of industrial capitalism, two realities helped propel the rise of environmental studies. First, the planet is, of course, finite and so is most of the resources we exploit, so our current system is not sustainable. And second, the rapid rise in resource exploitation driven by industrial capitalism causes equally rapidly rising environmental residue in the form of air, water, and soil pollution and imperils ecosystems all over the globe.

Conservation

Out of the humanist impulse came the first widespread conservation movement, emerging in the United States in the late 1800s. *Conservation* brings together ideas from economics, such as scarcity, with those of agriculture and animal management, such as the need for predictable feed stocks. When resources become scarce, then conserving what is left becomes a logical and ethical priority. For the conservationist, one need not wait until the point of scarcity; rather, conservation is seen an ethical way to live no matter how many resources remain. The idea is most commonly embodied by Gifford Pinchot (1865–1946), the first head of the U.S. Forest Service, who famously made decisions based on the greatest good for the greatest number of people. The conservationist sees the need to use our natural environment to improve the human condition, but those benefits should be made accessible to the largest possible number of people and with minimal waste, so that there will continue to be resources for use in the future. This is considered by many to embody the idea of enlightened self-interest. If natural resources get squandered, many of them cannot be replaced, so thinking in terms

of long-term sustainability of the resource makes more logical sense than exploiting as much as possible now, for short-term gain. Although almost nobody disagrees with the underlying concept that conservation makes sense, exactly where to draw the line between responsible use and overexploitation of any resource comes down to shifting social priorities and tensions related to who gains from resource exploitation and who does not.

Preservation

Going farther than conservation, *preservation* as a philosophy or mode of thinking promotes the idea that some things in nature must be preserved and never used by humans, on the grounds that their existence adds to the aesthetic beauty, the grandeur, stability, and power of nature. Natural systems must be protected from abuse, by preserving them from any exploitation at all. There is no middle ground for preservationists, perhaps the most famous of whom was John Muir (1838–1914), who fought a lengthy political battle of wills with Gifford Pinchot over whether the Tuolumne River in California should be dammed to provide drinking water for San Francisco after the 1904 earthquake. To dam the river would mean flooding one of the most exquisite valleys in America, and Muir believed passionately that the valley's intrinsic beauty and value could never be compensated for once it was lost under the rising water of a reservoir. In the end, it is arguable that Muir lost the battle but won the war. While the Hetch Hetchy dam did get built, the political explosion and fallout that followed set off the first widespread efforts at environmental preservation in the United States. One remnant of that fight remains in the form of the Sierra Club, a nonprofit organization that is now nationwide, started by John Muir at the turn of the last century. The concept of preservation now has a persistent voice in our conversations about environmental issues, even if there may be a lack of clarity on just what or how much nature should be preserved.

Tragedy of the Commons

In 1968, Garrett Hardin, an ecologist, set off a storm of debate across several fields of study with an article in the journal *Science* titled “The

Tragedy of the Commons,” in which he reconsidered the social question of population issues, but with a focus on economic theory and natural resources. To the present day, his argument continues to be a theoretical touchstone for much of the environmental and natural resource communities, while its implications echo outward to many disciplines. According to Hardin, when individuals have unlimited access to natural resources, they will inevitably act in their best interests to maximize their own utility; that is to say, maximize their ability to *use* the raw materials of the environment to satisfy needs or wants. If each individual profits by exploiting as much of the resource as possible, but the group as a whole shares in the negative utility caused by depletion of the resource, then each individual would be compelled to overuse the resource, since the individual gains all of the profit but shares only a fraction of the impact felt from the degraded resource.

In his article, Hardin illustrated his point using the example of common grazing lands for herders. Each individual will seek to maximize his profit from the resource, so it would be in their best interest to add one more sheep to their herd. Every herd owner does the same thing, and the logic appears sound. Each individual herd owner increases their own herd, thus their profits, and every individual gains wealth. The dilemma comes when the common resource, in this case the common grazing area, becomes degraded to the point that it can no longer be used at all, and the grazing economy collapses. In fact, the herders are locked into a system that compels each individual to grab as much of the resource as possible at ever increasing rates. In the case of the herders, if an individual does *not* add more sheep to their herd, then other competing individuals will simply add more and the herder will miss the opportunity to maximize utility. The hard-hitting message contained in the example rests with the realization that it is possible for each individual to gain by acting in their own self-interest, while the group as a whole loses in the end.

This exposes a dangerous but previously unaddressed problem at the heart of capitalism and the neoliberal economy. One of the fundamental tenets of neoliberal economy assumes that all actors seek to maximize their own utility and

should behave accordingly, and this is the basis of rational actions. But, if all actors behave rationally, then they will seek to maximize their utility in the market, even as the resources upon which the economy depends become degraded then depleted, and eventually disappear completely. This leads to total economic and ecologic collapse, precipitated by the destruction of our natural environment. Hardin applied this logic to the threat of overpopulation, viewing the planetary carrying capacity as the commons, and each new person to feed as a new drain on the common resource of the planet, and concluded that there is no technical solution to this dilemma. In his analysis, Hardin separated out those problems which might have a technical solution, or a change in the techniques or science applied to them, from other problems for which there is no technical solution; this latter category requires a change in our morality or human values to find a solution. From this dialogue, efforts to clarify and express just what the range of human values were resulted in the widespread use and discussion of the terms anthropocentrism, ethnocentrism, and ecocentrism. Clearly anthropocentric values should lead economic actors to maximize their own profit in the short term, even at the cost of destroying our natural environment over time and would spiral toward a tragedy of the commons without pause.

Critics of Hardin, most famously Elinor Ostrom, point to counterexamples, where a tragedy of the commons was avoided, and suggest that it is not as prevalent or as inevitable as Hardin would have us believe. She and other critics point to examples in which cooperative behavior results in collective restraint and sustained use of resources, since it is understood that both individuals and the collective benefit from this action. Thus, if economic actors share ecocentric values, then they would cooperate enough to collectively restrain the total destruction of the environment, thereby providing the resources they share, need, and can sustainably use. Others criticize Hardin for being historically inaccurate by failing to acknowledge the demographic transition or to distinguish between common-pool resources and open-access resources.

The *tragedy of the commons* metaphor aptly illustrates that unlimited demand in a world of finite resources leads to overexploitation and

ruination of our environment—given open access to common resources. *The tragedy of the commons* theory sparked debate in circles of thought as wide and seemingly unconnected as game theory, psychology, cultural anthropology, and strategic model building, as well as its more obvious links to environmental studies, geography, political science, ecology, and economics. The theory stands as an easily graspable conceptual model for a wide array of resource-related and environmental issues based on the dilemma of finite resources such as water, forests, fossil fuels, and arable land, for which we seem to have an ever increasing number of users. Environmental scholars and activists remain staunch, however, in their cries for vigilance, and for the need to rework our priorities and our mechanisms of constraint. They believe that the scale and persistence of environmental degradation speaks for itself loudly, pointing in no uncertain terms toward global tragedy and environmental calamity if the pace and manner of global development, and the values of endless consumerism, are not curtailed—and quickly.

Gaia Theory

Developed by the chemist, physician, and environmentalist James Lovelock and the microbiologist Lynn Margulis in the 1970s, *Gaia theory* suggests that the biotic elements of Earth interact with inorganic elements of our environment to form a self-regulating interconnected system that contributes to maintaining the conditions for life on the planet. Named after the Greek goddess Gaia, this concept, also known as the *Gaia principle* and *Gaia hypothesis*, suggests that organisms coevolve with their environment. In other words, biota influence aspects of their inorganic environment, such as temperature or atmospheric conditions, and that altered environment in turn influences the development of biota; together, these forces generate a stable system, described as *planetary homeostasis*. The interactions are complex, but here are two simplified examples used in Gaia theory: air temperature and atmosphere. Since the beginning of life on planet Earth, there has been a 25% increase in the amount of heat from the sun, but the surface temperature has remained virtually constant. Also, our atmosphere consists of roughly 79% nitrogen, 20.8% oxygen,

and 0.03% carbon dioxide (CO₂). This highly unstable mix of reactive gases could not be maintained without constant replacement and/or removal by the Earth's biota. Earlier in Earth's history, the atmosphere contained higher levels of CO₂ and these CO₂ levels produced a greenhouse effect that warmed the planet. Later, biota on the planet has produced increasing amounts of oxygen and methane, via cycling through oceans and rocks; these changes in the atmospheric chemistry have been maintained by evolving species. Now, CO₂ is minimized by pumping down of carbon through the biota. Without the presence of greenhouse gases in the atmosphere, scientists estimate that our planet's surface temperature would be approximately 19°C (66.2°F).

The idea of a planetary homeostasis influenced by biotic factors already exists and is studied by scholars from academic disciplines including ecology and biogeochemistry, or Earth system sciences. However, a unique element of Gaia theory puts forth the assessment that this homeostatic balance is pursued actively, if unconsciously, to achieve optimum conditions for life on the planet—even if external forces threaten that stability. Thus, taken to its logical conclusion, the theory implies that the biotic balance with the regulatory systems of the planet, such as oxygen in the atmosphere or salinity in the ocean, will maintain balances needed to optimize a biogeochemical balance, even if external human activity threatens that equilibrium. Gaia theorists believe that humans could find themselves one day headed toward extinction as a species but that the planet as a whole would still remain, with its regulatory systems intact. Gaian theory essentially posits that the Earth is in some complex sense *alive* and functions as an intricately interconnected singular entity.

This theory has generated much speculation and debate along the academic spectrum, from philosophers to Darwinists. Philosophers use Gaia theory to begin asking fundamental questions about teleology, the meaning of *living*, and the need for consciousness in order to act in unison. To achieve this, reconsiderations of basic ethical values or concern for nonhuman and even nonliving components of the natural world would be required. The Gaian worldview rests on the notion of large-scale cooperation rather than competition as a driving force and simultaneously suggests that

the planet's natural systems are not just interconnected but also function toward a common purpose—stabilization. On the other end of the spectrum, staunch Darwinists critique the Gaia theory and suggest that evolutionary strategies could never account for complex interaction among nonsentient biota at such large scales. For Darwinists, there is simply no plausible way that the feedback loops Lovelock says work to stabilize the planet could have evolved. For Darwinists, some mechanism must be shown to control the processes described, and if they see no mechanism at work, they see no larger cohesive end. Gaia theory has also set off arguments among scientists about what it means to be *living*, and Darwinist critics note that since Gaia cannot reproduce itself, it fails to meet one of the key components in the accepted definition of *living* among the scientific community.

Sustainability

This concept is linked to, but not the same as, the Gaia theory. The term *sustainability*, a recent addition to our lexicon, has become so widely used and with different definitions that the term may have lost some of its original power as an idea. The most prevalent definitions all agree that in order to for the way we live to be sustainable, we must treat our natural environment in ways that ensure we have enough resources left that future generations are not significantly disadvantaged. Although there may not be a consensus on what living *sustainability* entails, there is agreement on what it is not. Overexploitation of resources and rampant environmental degradation are signs that we are not living within our ecologic means. Capitalist economies are based on individual incentive to maximize profit, which necessitates exploiting more and more resources, which generates ever more pollution. By predicating economies and policies on nonrenewable fossil fuels, we are essentially spending the environmental bank accounts of future generations to satisfy this generation's lifestyle, which is clearly unsustainable.

Pathways toward achieving sustainability, as currently defined, have become vaguer, more varied, and more contentious. Some critics of the status quo suggest that we can put off choices regarding whether or how to include future

generations in our decisions by implementing simple conservation techniques. If priorities changed at every scale from individuals to federal governments to include conservation, thus reducing the detrimental aspects of capitalism, we could delay reaching the limits of fossil-fuel-based economies for several more generations as we think constructively about what to do.

Carrying Capacity

Borrowing a classic concept from biology, practitioners of environmental studies continually grapple with the quantitative and qualitative metrics of *carrying capacity*, or the amount of population that an ecosystem can sustainably support without overtaxing the resources and balance of life within the ecosystem itself. Any discussion of sustainability must be framed by asking pertinent questions about just how many people can be sustained and at what scale, or at what level of existence. The problem rests with the reality that we simply do not know what the biologic carrying capacity of the planet is, and we are not likely to know until we pass the threshold which the planet can support, then see the crash in population due to our having reached this tipping point and not stopped. This is not a number or proportion that can be easily estimated, and even if we could do so, that begs the question of the standard at which the population would live. The carrying capacity would be quite different if everyone were at the same standard of living as people in Guatemala, for example. And it is not even possible today to have the entire planet live the same lifestyle as residents of the United States, which contains less than 5% of the world's population but consumes well over 25% of all its resources each year.

Even without knowing the precise number of people this planet might sustain, many observers base their critiques of the status quo on maldistribution of resources and vast inequities in wealth and poverty across the globe. They recognize that questions of sustainability must address *how* people live, not simply *how many* people live on the planet. A formula borrowed from geographers illustrates this idea: $I = P \times A \times T$ estimates total Impact by multiplying Population times Affluence and Technology. For instance, if a country is quite affluent, then it will probably have access to much

technology, and the overall impact of each person will be far greater than the relative impact of a single person elsewhere. A new baby born in the United States last year is expected to consume 25 times more resources than a new baby born in Burma, based on this equation. Thus, small but affluent populations have total environmental impacts much larger than their population alone would suggest. Similarly, very populous but poor countries without access to technology languish at the bare edge of subsistence using far fewer resources than their numbers alone would predict.

This formula does not account for everything, and critics are quick to point out counterexamples, which do not fit neatly into this formula. How does $I = P \times A \times T$ deal with countries that utilize technology to develop renewable resources, for instance? Another situation not easily accounted for is that of a country that uses technology, but that country's population does not acquire affluence associated with the level of technology. This lopsided situation occurs in underdeveloped nations with large labor forces working for low wages in outsourced manufacturing jobs that produce items consumed by developed countries.

Environmental Justice

A new and rising trend that environmental studies scholars have helped bring to the forefront of conversation is *environmental justice*. The idea of environmental justice centers on the imbalance between those who benefit from using resources versus those who live with the residue of that use, or the environmental degradation generated as a byproduct of exploiting, refining, distributing, and consuming natural resources at all phases. Thus, there is a need to find justice for those who disproportionately bear the burden of others' comfort. For instance, only those who are too poor to move elsewhere will live near city dumps, toxic brownfields, or the noise, smell, and filth associated with factories. There is also a racial component to environmental justice in the United States, due to the high numbers of minorities who are lower on the socioeconomic ladder and thus have fewer options to relocate when faced with pollution, toxic health hazards, or even an abundance of ugly industrial buildings in their proximity. Environmental justice advocates seek to redress these imbalances, often through the

legal system and politics, with varying amounts of success. One intriguing new direction in this rapidly evolving subdiscipline rests with geographic inaccessibility and the persistence of poverty. Wealthy people take the best land, with the best resources, leaving only marginal lands for others, which accelerates environmental degradation specifically because the areas were marginal to begin with, then become overexploited. Once the marginalized ecosystems lose their fragile equilibrium, then a cycle of degradation ensues, leading to overexploitation, which in turn creates environmental destruction, such as mudslides, desertification, and famine, which leads further to desperation, and this continues or even accelerates the overexploitation. Such is the situation in areas such as the slums of Karachi, Pakistan, and nearly the entire Ganges River delta region of Bangladesh.

Deep Ecology

Initially coined as a term by Norwegian philosopher Arne Naess in 1973, *deep ecology* draws together ideas found in ecocentrism, Gaia theory, and environmentalism. Naess claims that deep ecology is merely the philosophical extension of the scientific foundations of ecology. Essentially, humankind is not above nature but merely one part of the larger ecologic system that is the living Earth. At its core, deep ecology stresses wilderness preservation, human population growth limits, and the need to *live simply* on the planet or, in other words, to leave a light ecological footprint. Deep ecologists contend that unbridled industrial capitalism in itself is the major problem escalating exploitation and destruction of the natural world, citing rapid species extinction, reduced biodiversity, and accelerating climate change; they believe the only moral response to our conundrum lies in fundamentally changing the way we structure our political economies. Deep ecology can be seen as the core component of larger movements such as the rise of *green party* politics, and contemporary *eco-spirituality*, which runs counter to the anthropocentric Judeo-Christian traditions. In deep ecology, we see a philosophic repositioning of the concept of *self* as well as the wider metaphysical implications of humans as just one piece of a single but interconnected unfolding reality.

Some critics see deep ecology as a philosophy that denigrates or dislikes humanity, since the central tenet revolves around ecosystem health and humans as just one piece of this system. Needless to say, deep ecology presents the most significant challenge to the dominance of industrial capitalism, anthropocentrism, resource exploitation, and large governments since the beginning of capitalism.

Gina Bloodworth

See also Climate Science; Natural Resources

Further Readings

- Hardin, G. (1968). Tragedy of the commons. *Science*, 162, 1243–1248.
- Ostrom, E. (1999). Revisiting the commons: Local lessons, global challenges. *Science*, 284(5412), 278–282. <https://doi.org/10.1126/science.284.5412.278>
- Lovelock, J. (1995). *Gaia: A new look at life on earth*. Oxford University Press.
- Naess, A. The shallow and the deep, long-range ecology movement: A summary. *Inquiry*, 16, 95–100.

EPISTEMOLOGY

Epistemology is a branch of philosophy that explores knowledge or how people know what they know. Epistemological questions include “What counts as knowledge?” “How is knowledge acquired?” “To what extent is knowledge provable?” and “What is the relationship between the knower and what is known?” Ways of knowing have been contemplated and contextualized broadly across many disciplines through discussions related to reason, emotion, perception, and language and, in more specific terms, via discussions about the method of tenacity (e.g., superstition), method of authority (e.g., religion), method of intuition (i.e., philosophy), and method of science. The *method of tenacity* believes something is truthful despite evidence. The *method of authority* is accepting a truth to be true because someone said it was so. The *method of intuition* is accepting a truth to be true because it is rooted in logical reasoning, yet no empirical support exists. The *method of science* seeks to find truth through observable events.

Epistemological questions about knowledge and theory building are often accompanied by three related philosophical inquiries and/or claims: ontological, praxeological, and axiological. *Ontology*, or the study of reality, is interested in answering questions such as “What can be said to exist?” and “What are the meanings of beings?” *Praxeology*, or the study of human action, is interested in questions such as “What is the practice of the practical argument?” *Axiology* makes claims or asks questions related to the worth of knowledge or theory by asking questions such as “What is the role or function of scholarship in society?” Additionally, methodological questions/concerns, or philosophical claims about what methods are appropriate and/or effective given the theory or inquiry, must be addressed.

This entry comprises seven main topics that help to define and contextualize epistemology. First is an overview of its history and the role that epistemology plays in theoretical inquiries. Then, a division of epistemology into two main camps, *internalism* and *externalism*, is discussed followed by another main division, here of epistemological accounts of the structure of knowledge as either *foundationalist* or *coherentist*. Following these stage-setting discussions, consideration is given first to how epistemological concerns shape and are shaped by the research paradigms of inquirers, second the meta-epistemological orientation of objectivism and subjectivism, third the role of epistemology in self-reflection, and finally of the main tenets of feminist epistemology.

The Historical and Theoretical Role of Epistemology

The word *epistemology* derives from the Greek language: *episteme* (know) and *logos* (logic). The philosophical study of epistemology dates back to antiquity when the Sophists questioned the possibility of objective truth, Plato argued that absolute or exact knowledge existed, and Aristotle held that experience was the key to knowing. The introduction of René Descartes’s (1596–1650) book, *Meditations*, in 1641 demarcated a shift away from Aristotelean ways of knowing by introducing a more modern philosophy centered on skepticism or the method of doubt. In the late 1700s, Immanuel Kant (1724–1804) introduced

a newer philosophical shift in epistemology by challenging both empiricism and rationalism. His critical work deconstructed the two competing belief structures of the Enlightenment era: Knowledge was acquired through either reasoning (a priori knowledge; rationalism) or sense perception (a posteriori knowledge; empiricism). Beginning in the 19th century, as scholars attempted to reconcile rationalism and empiricism, discussions about how each approach or understanding of knowledge creation should address questions of justification and description emerged. The debates about justification, in particular, are covered in more detail in the following section of this entry.

Epistemological questions about the nature, scope, and function of knowledge are primary to the development of theory. In fact, it is an essential precursor to theoretical perspectives, methodological considerations, and method choices across many social scientific and humanistic academic disciplines. The *tripartite theory of knowledge* suggests that knowledge is a “justified true belief” More specifically, belief, truth, and justification are all necessary factors of knowledge—and of theory building. That is, unless one believes something, they cannot know it; if one knows a thing, then it must be true; and in order to know a thing, one must have a good reason for doing so (it is not enough to believe it is true). Thus, theories about human phenomena (or any phenomenon) are rooted in the basic desire to explain, understand, predict, and/or control individual and/or shared realities through belief, truth, and justification.

Internalism and Externalism

The internalism–externalism debate lies near the center of modern discussions about epistemology. The basic idea of internalism is that justification is solely determined by factors that are *internal* to an individual. Externalists disagree, asserting that justification depends on additional factors that are *external* to an individual. The rise of this debate coincides with the rebirth of epistemology after Edmund Gettier’s famous 1963 paper, “Is Justified True Belief Knowledge?” In that paper, Gettier presented several cases to show that knowledge is not identical to justified true belief. Cases of this type are typically referred to as *Gettier cases* and

illustrate *Gettier problems*. Traditionally, Gettier cases show that one can have internally adequate justification without knowledge. The introduction of the Gettier problem to epistemology required scholars to rethink the connection between true belief and knowledge, and the subsequent discussion generated what became the internalism–externalism debate over the nature of justification in an account of knowledge. Internalists maintained that knowledge requires justification and that the nature of this justification is completely determined by a subject’s internal states and/or reasons. Externalists argued that either the nature of justification is not completely determined by internal factors alone or knowledge does not require justification. The internalism–externalism discussion engages a wide range of epistemological issues involving the nature of rationality, the ethics of belief, and skepticism.

The Structure of Knowledge: Foundationalism, Coherentism, and Reliabilism

The three most prominent theories of epistemic justification are *foundationalism*, *coherentism*, and *reliabilism*. All three are meant to address the problem known as the *regress problem of justification*. More clearly, if individuals think of epistemic justification in inferential terms (i.e., for every justified belief there must be at least one other justified belief on which it is based, which must in turn be based on at least one other justified belief, etc.), then there must be an infinite regress of justified beliefs—and individuals do not seem to have the ability to make sense of an infinite chain of inferences.

Foundationalism is an attempt to stop this regress of justification by positing the existence of foundational or *basic* beliefs (i.e., justified without being inferred from other beliefs). Because basic beliefs are justified, they are able to confer justification onto other beliefs that can be inferred from them. Coherentism is a rival theory of justification. Unlike foundationalists, coherentists reject the idea that individual beliefs are justified by being inferred from other beliefs. Instead, according to coherentism, whole systems of beliefs are justified by their coherence or the idea that a belief set is coherent to the extent that it is

consistent, cohesive, and comprehensive. Reliabilism is an alternative theory of justification to foundationalism and coherentism. According to reliabilism, a belief is justified based mainly on how it is formed—thereby removing the burden of deciding whether a belief is justified based on whether or not it is appropriately related to other beliefs. Essentially, the reliabilist holds that a belief’s justification depends on whether it is formed using a reliable or an unreliable method (e.g., if perception is a reliable method for forming beliefs, then beliefs based on perception are justified).

Research Paradigms

In the social sciences and humanities, epistemological concerns become paramount depending on the paradigm the researchers are positioned within. In the social sciences, *postpositivism*, *interpretivism*, and *critical theory* comprise the primary paradigms. The postpositivist paradigm, largely informed by the French philosopher Auguste Comte (1798–1857), emphasizes observation and reason as means of understanding human phenomena. This paradigm is anchored by the underlying assumptions of determinism, empiricism, and generality—and the belief that true knowledge may be obtained by observation and experiment. *Determinism* means that events are caused by other circumstances; and hence, understanding such correlatory and casual relationships is necessary for prediction and control. *Empiricism* aims to collect verifiable empirical data in support of theories or hypotheses. *Generality* is the process of generalizing the observation of the particular phenomenon to the larger population. The postpositivist paradigm systematizes the knowledge-generation process with the help of quantification via survey, content analyses, and experiments.

Interpretivism emphasizes that social reality and truth/reality are multilayered and complex, and a single phenomenon is likely to have multiple interpretations that may only be understood by probing the various unexplored dimensions of a phenomenon rather than establishing specific relationships among variables. It is grounded in three major areas of study: *phenomenology*, *ethnomethodology*, and *symbolic interactionism*. All

three emphasize human interaction with phenomena in one's daily life and suggest qualitative rather than quantitative methodological choices. *Phenomenology* is a theoretical perspective that suggests that individual behavior is determined by the experience of an individual's direct interaction with a phenomenon. During this interaction, individuals attach meanings to different actions and or ideas and thereby construct new experiences. *Ethnomethodology* deals with the world of everyday life. According to ethnomethodologists, theoretical concerns center on the process by which commonsense reality is constructed in everyday face-to-face interactions. This approach studies the process by which people invoke certain *taken-for-granted* rules about human behavior and give it meaning. Symbolic interactionism focuses on trying to understand and interpret interactions that take place between individuals. The peculiarity of this approach is that human beings *interpret and define* each other's actions instead of merely *reacting* to each other's actions. Human interaction in the social world is mediated by the use of symbols like those of language, which help human beings to give meaning to objects. To understand and explain phenomena, the interpretivist paradigm uses ethnography, case studies, and biographical data collection as its primary methods.

Critical theory is a tradition that assumes reality is created and shaped by social, political, cultural, economic, ethnic, and gender-based forces that have been reified over time into social structures that are taken to be natural or real. Additionally, critical theorists state that we as individuals, including researchers, cannot separate ourselves from what we know and this inevitably influences inquiry. What can be known is inextricably tied to the interaction between a particular investigator and a particular object or group. Critical theory is particularly concerned with the role of power in all interactions—and that the perpetuation of the subjective—objective controversy is problematic. Rather, critical theorists suggest that *objective* practices are often the most *subjective*. Methodological approaches tend to rely on dialogic methods—that is, methods that combine interviewing and observation with approaches that foster conversation and reflection. This reflective dialogic allows the researcher

and the participants to question the *natural* state and challenge underlying and often systemic power structures.

Objectivist and Subjectivist Positions

Because all theories are based in metatheoretical assumptions, it is important to understand two key epistemological positions. First, the objectivist position, which was dominant in physical and social science traditions throughout the 20th century, espouses empiricist and rationalist ideas. Objectivists generally believe that reality exists outside of and apart from the individual. From this perspective, a knowable reality awaits discovery through the scientific method. As such, objectivists explain social phenomena almost exclusively through causal relationships. By contrast, the subjectivist epistemological position assumes that human beings actively create knowledge. From this frame, because knowledge is cocreated in social interaction, it is situated and highly contextual. Subjectivists favor interpretive processes when developing theory. Rather than seeking to explain social phenomena through universal laws, subjectivists aim to describe and understand various contexts wherein knowers and the known function. The objectivist and subjectivist positions serve different interests, and each of these positions differs significantly from a third epistemological stance that considers the political nature of knowledge.

Self-Reflection as Epistemological Project

Renowned critical theorist Jürgen Habermas outlines a *politics of epistemology* (Mumby, 2001) to highlight the relationship between power and knowledge. Demonstrating that scientific knowledge is but one type of knowledge in the world, Habermas identifies three *knowledge constitutive* interests in society—the first two of which align with the objectivist and subjectivist positions, respectively: (1) *the empirical-analytical cognitive interest*, rooted in technical goals to predict and control both the physical and social world; (2) *the hermeneutic-historical cognitive interest*, rooted in goals to understand situated human activity; and (3) *the critical-emancipatory cognitive interest*, which posits knowledge as a

self-reflexive process through which to uncover hidden but pervasive power that shapes taken-for-granted assumptions about reality. Critical theorists thus differ from objectivists and subjectivists in that their epistemological commitments view knowledge as a means for transformation and emancipation. Habermas (1971, p. 197) explains, "In self-reflection, knowledge for the sake of knowledge comes to coincide with the interest in autonomy and responsibility." For critical theorists, knowledge is political. That is, knowledge (and the ways it is created and expanded) is intrinsically linked to ideology that can function as a form of control. Antonio Gramsci (1891–1937), Marxist theoretician, stated that this form of control is also known as *hegemony*, a process in which dominant individuals and groups lead others to accept and unconsciously exercise subordination as natural. While there are many varieties of critical social science, critical theorists are united in their goals of providing normative theories as the basis for social inquiries and effecting social change.

Feminist Epistemology

A prominent and increasingly influential area within the critical tradition is feminist studies. Feminist studies, too, encompass a wide range of viewpoints. In general, theorists working in feminist studies share an approach to epistemology that recognizes the connection between the knower and the known. Standpoint feminists, for example, argue that all women do not speak with a single voice, and they caution researchers against universalizing the experiences of women. Feminist theorists overwhelmingly reject knowledge that is grounded in totalizing theory, arguing instead that knowledge about the social world is neither provable nor objective. As feminist communication scholar Cindy Griffin notes, theory is a subjective compilation of explanations that are bound by the ideologies of the researcher and deeply influenced by the social milieu from which they emerged.

Courtney Vail Fletcher and Rachel Stohr

See also Communication Theory, Technical Overview; Knowledge; Philosophy of Science; Theory Construction; Truth

Further Readings

- Alston, W. (1986). Internalism and externalism in epistemology. *Philosophical Topics*, 14, 179–221.
- Gettier, E. (1963). Is justified true belief knowledge? *Analysis*, 23(6), 121–123.
- Gramsci, A. (1971). *Selections from the prison notebooks* (Q. Hoare & G. Nowell Smith, Trans.). International Publishers.
- Habermas, J. (1971). *Knowledge and human interests* (J. Shapiro, Trans.). Beacon Press.
- Mumby, D. K. (2001). Power and politics. In F. M. Jablin & L. L. Putnam (Eds.), *The new handbook of organizational communication: Advances in theory, research, and methods* (pp. 585–623). Sage.

EVIDENCE

Evidence is the information currently available, at a given time, to support or disprove an assertion, notion, or theory. In the medical, educational, and social services realms, there has been an increasing emphasis on practices that are supported by evidence. Unfortunately, there is often confusion about what constitutes solid evidence, and many people use tenuous evidence to support their assertions and programs. Using weak evidence to inform decisions that affect many people's lives is irresponsible and could cause unnecessary harm while also wasting precious financial resources. Evidence from scientific studies is usually considered the most rigorous, but it is important to understand the different levels of evidence that can be created from scientific research. When complex theories are involved, there is often not enough current evidence to prove or disprove them, setting the stage for new scientific studies. With more data available now than at any time in history, it is important to discern the continuum between weak and strong evidence in order to make wise, evidence-based decisions and create effective programs to increase productivity, health, achievement, and well-being. In the following sections, this entry traces the increase in evidence-based practices over the past half-century and explores the nature of evidence itself. Among the questions addressed are what constitutes weak versus strong evidence for various claims, what types and levels of evidence constitute the

strongest support for scientific theories and hypotheses, which types of experimental designs are capable of producing the most reliable scientific evidence, and what sorts of instruments are most capable of assessing the value of purported evidence for the efficacy of various practices in, for example, education. The distinction is made between a lack of evidence to support a claim and the existence of evidence showing that a given product, program, or hypothesis is ineffective or harmful. The entry concludes with a discussion on the importance of evaluating evidence for the benefit of society and its institutions and for the sake of personal well-being.

Increased Emphasis on Scientific Evidence

The demands for evidence-based practices have increased since the mid-20th century for medicine, education, psychology practice, and in various other fields. In other words, administrators and practitioners are being held more accountable to demonstrate that their practices have a positive effect. For instance, the federal government now strongly encourages public schools to use instructional techniques that are supported by scientific evidence. Furthermore, schools across the United States must report on their students' progress each year, often measured by performance on statewide achievement tests. Unfortunately, many of the states' measures of achievements have not been adequately validated in peer-reviewed scientific journals, suggesting that achievement gains on state achievement tests may not be reliable and valid evidence of positive educational effects.

In the health care field, the pressure to provide evidence for the effectiveness of practices comes in part from the U.S. Food and Drug Administration (FDA) when it comes to medication. Medications are expected to demonstrate effectiveness through significant improvements in the health of a group of patients who receive the medication (the treatment group) compared with a group of patients on a wait-list or receiving a placebo (e.g., a sugar pill) over the course of a few weeks or months. Furthermore, the medication should not produce an inordinate frequency of dangerous side effects in the treatment group. In addition to the FDA, insurance companies may provide the impetus for using evidence-based practices by having a list of

approved procedures (or procedures they fund at a higher rate) that are supported by adequate scientific research.

Increasingly, private foundations and providers of federal grants emphasize evidence in determining what they will fund. Often, grant writers and investigators must make a clear and convincing argument that the components of the program they wish to initiate are informed by reputable theories and are supported by rigorous evidence. Furthermore, they must often explain how they will evaluate whether grant-funded programs will have the intended effect in the hospital, community, city, or state in which they will be applied. Although such program evaluation is not usually as rigorous as peer-reviewed scientific research (i.e., research reviewed by a group of expert scholars or scientists who do not know the author's identity), it is expected to be based on strong evidence.

Weak Evidence

Not all evidence is created equal. Many people assume that they have evidence for an assumption, folk theory, or pattern in life because they have seen it happen a few times in their own life or in the lives of people they know. For instance, let us say that Charlie feels compelled to wear a jester's hat every Sunday because his favorite football team has won most of their games when he wore the hat. He has even made self-deprecating remarks when he supposedly caused his team to lose by his not wearing the hat. Most people can tell that his wearing a jester's hat while he watches the game at a local bar has no effect on the outcome of a professional football contest, but Charlie swore by it until his team finally lost four games in a row. This type of superstitious thinking is an example of basing beliefs on flimsy evidence. It is also an example of what is known as *illusory correlation*, whereby two things seem related whereas other, more important factors (such as the injury list for the favorite football team and the records of the opposing teams) are ignored.

As a somewhat more insidious example of forming and conveying beliefs based on limited evidence, let us say that a dedicated teacher has seen three highly energetic African American boys from single-parent homes struggle academically,

and their performance only seemed to improve when later receiving special education classes or medication. If not rigorously and objectively analyzing the immediate and wider evidence, this teacher might assume that highly energetic young African American boys from single-parent homes who struggle academically are at a high risk for attention deficit hyperactivity disorder (ADHD) and only improve academically with special education or medication. Numerous research studies indicate, however, that many students do not improve academically in special education and that many students with ADHD continue to struggle academically even after receiving medication. There have also been studies indicating that African American boys are over-referred by teachers for ADHD-related concerns, even when girls and non-African American boys display similar levels of inattention, impulsivity, and hyperactivity during observations by an independent observer (a highly trained observer who has no personal connection to the people being observed). A keen clinical psychologist, school psychologist, or mental health consultant will be aware of this scientific evidence and help the team of educators analyze all existing data and then form *hypotheses*, which are clear and testable statements about what is likely causing the problem. Based on these hypotheses, the team would then develop interventions that strategically target the child's specific constellation of problems while also closely measuring the students' response to the interventions. Often, this leads to improved behavior and academic performance without special education or medication, but in other cases, the teacher's original impressions concerning the particular child may be proven correct. The important thing is that well-informed hypotheses are rigorously tested, so that professionals do not make major decisions that affect people's lives based on assumptions and limited evidence. Rather, rigorously testing hypotheses leads to richer evidence, making it possible for professionals to make decisions in light of stronger evidence.

The weakest level of evidence is that which relies either upon one's limited experience alone or upon the opinions of others who are basing their assertions on their own limited experience and knowledge. Scientists endeavor to avoid this by conducting studies involving many participants,

usually with whom they have no personal relationship, thereby promoting objectivity (seeing things clearly with much less bias). Furthermore, before scientists form hypotheses, they will usually diligently read the scientific literature on the topic, helping them to ground their views in relevant information and related theory rather than in the opinions that they and others might have expressed in person or in personal correspondence or popular media (e.g., blogs, newspapers, or magazines).

Levels of Scientific Evidence

Another generally weak form of evidence is that of relying heavily on percentages from surveys. Surveys, when they sample people properly, let us know what percentage of the population in a company, school, state, or nation view things in a certain way, but they tell us little about why. Furthermore, many surveys use what is called a *convenience sample* (e.g., the person at the mall asking you for "just 15 minutes" to fill out their survey on a clipboard or smartphone). The problem with convenience samples is that they usually do not represent the population of interest reliably. In the case of the mall survey, people with more time, people who are less assertive, or people who are very agreeable might be more likely to fill out the survey, providing skewed results (i.e., results that do not represent the whole group of people that the surveyor hopes to assess). Better survey studies use a random sample, such that everyone within the population being studied has an equal chance of being included in the study, yet survey studies are still problematic unless they address control variables, which are other factors that may affect results, such as the socioeconomic status (a mix of education level, wealth, and current occupation), gender, and ethnicity of the person filling out the survey.

Statistical correlations are an important form of evidence, but they are generally considered preliminary data. A *correlation* is the strength of the relationship between two variables. For instance, if there is a statistically significant correlation between how often parents read to their children (i.e., shared reading frequency) and how much those children achieve, this is interesting. However, this correlation could be illusory in that some

other variable related to both of these variables could explain the relationship. For instance, children who achieve more may have highly educated parents who model a fascination with learning more than most parents, which could lead to higher achievement. So, if we want to know whether shared reading frequency is causing improvements in achievement, we would seek higher levels of evidence. One might start by determining whether shared reading frequency predicts future achievement even when controlling for the effects of previous achievement, gender (girls often have slightly higher reading achievement), race/ethnicity, and family socioeconomic status. If these results are as predicted, one might proceed to conduct an even more rigorous study by randomly assigning students to either a treatment group that receives a parent-child shared reading intervention or a control group that is on a wait-list to receive the intervention later. If the treatment group improves significantly more than the control group, as demonstrated by statistical tests, we gain more confidence that parent-child shared reading causes an improvement in reading achievement. The evidence for the effectiveness of parent-child shared reading (or any other intervention) becomes even stronger when the findings are replicated in subsequent studies, especially when at least one of the subsequent studies is conducted by a research team that is independent of the developers of the intervention. This progression of research and correspondingly stronger evidence has been actualized in the realm of parent-child shared reading frequency. Even though scientists are rigorously trained to be objective, results replicated by another research team help to show that others can apply the methods successfully as well. Furthermore, the independent research team may have less interest in the intervention and thereby be less susceptible to introducing bias into the experiment, although scientists have ways of reducing bias, such as by having the data collected and entered by research assistants who are unaware of the hypotheses.

In the preceding example, families were randomly assigned to either a control group or an experimental group. Positive results from a randomized controlled trial (RCT) are usually considered the highest form of evidence in science, especially if the findings are replicated multiple

times with an adequate number of participants in each study. It is important that RCTs have enough participants because randomization disperses important unmeasured characteristics (e.g., motivation to improve, intelligence, and personality) between treatment and control groups increasingly evenly as the number of participants in a study increases. However, it is not always easy to use random assignment. For instance, people or organizations that agree to be in either the treatment or control group sometimes drop out of the study when they figure out that they are in the control group. That is, some people or organizations only feign agreement to participate in the RCT study in hopes of being assigned directly to the treatment group in order to receive an experimental treatment that they expected to be beneficial. Thus, random assignment to the treatment or control group is not always feasible or efficient.

The next highest form of applied experiment is the *quasi-experimental design*, in which people are recruited to join either the treatment or the control group, without random assignment. The people in the treatment and control groups may differ in some important way, but this can often be controlled for if effective pretest and posttest measures are employed. For instance, if one is studying how an intervention designed to treat symptoms of depression affects job performance, controlling for the prior skills, socioeconomic status, gender, mental health, and personalities of employees will help rule out important differences between the treatment and control groups. If the treatment group shows a statistically significant greater reduction in depression symptoms than the control group, researchers have found decent evidence that the treatment caused the improvement. However, only in an RCT with an adequate number of participants can we be sure that there is no difference between the groups (such as motivation to improve or some other unmeasured characteristic) that would explain improvement in response to a treatment. For this reason, positive evidence from quasi-experimental studies provides an impetus for following up with more rigorous experimental studies of the same treatment or program.

Although experimental and quasi-experimental designs help to create high levels of evidence for or against a hypothesis or proposition, they are relatively rarely employed. All too often programs are

evaluated in corporations, schools, states, and nations with very little rigorous evidence. For instance, countless programs are evaluated every year without any sort of control group. The biggest problem with not including a control group is that the design of the study makes it impossible to say whether or not the program or treatment caused any significant improvement that's been detected. Significant improvement detected within a study entailing solely a treatment group could be due to a host of other factors, such as people maturing (e.g., children developing cognitive skills as a function of natural child development), being more aware of their behavior when they know they are being assessed, or unmeasured interventions in other environments (e.g., a community-based support program positively affects employees' neighborhoods during the same time a work-based program/treatment is introduced). Conversely, when a program evaluation lacks a control group, the treatment group or program participants could make no improvement and the team could decide erroneously that the treatment or program did not work. This weak evidence could be used to decide to eliminate a potentially helpful program. For instance, let us say that a preschool and elementary school decide to promote parent-child shared reading in the home from preschool to first grade. After 3 years, the school administrators and educators analyze parent-child shared reading frequency data and are disappointed to find that the parents did not increase the amount of time they read with their children over the course of the 3 years. However, in a longitudinal study (a study that follows people over a long time) conducted by developmental scientists from Purdue University, Douglas Powell, Seung-Hee Son, Nancy File and John Mark Froiland found that many parents decrease their shared reading frequency (and other forms of home-based parent involvement) from preschool to first grade, when there is no special home literacy intervention in place. Thus, if the preschool and elementary school would have had a control group, they may have found the same thing, in which case the evidence would show that the treatment group remained the same while the control group showed a statistically significant decrease in reading with children at home. In light of the control groups' decreased performance, the

parent-child shared reading program might be considered a success because home literacy supports child development in various ways, but one would never know this without measuring the control group.

Longitudinal studies are not usually as rigorous as experimental and quasi-experimental studies, but they are an important improvement over studies that only assess people or groups once. In the preceding example of the study at Purdue, natural patterns among parents were discerned by observing them over a long period of time. That study also found that parents naturally increase their engagement in community-based activities (attending museums, zoos, public libraries, sporting events, and aquariums) with their children between preschool and first grade, which is important to know for anyone who wants to provide sound evidence that their program to increase community-based parent involvement has had a significant effect. Namely, program developers would need to demonstrate that their program had a positive effect above and beyond parents' natural development in this area. Longitudinal studies are also an improvement over studies with one wave of data because they establish what is called *temporal precedence*. Temporal precedence entails one variable occurring prior to another variable (e.g., high intrinsic motivation to help others and learn, predicting more success at work 2 years later). This helps to move toward establishing causation, but alone is not sufficient, which is why quasi-experimental or experimental studies are needed later. Longitudinal studies can help identify important patterns and help researchers, administrators, and policy makers to know whether the hypothesis is worthy of studying in a rigorous, sometimes risky, and often expensive experiment. In other words, it is important to gain as much knowledge and evidence as possible before conducting an experiment that may require more money and affect many people's lives in various ways.

The Importance of Assessment Tools or Outcomes as Contributors to Evidence

For any evidence involving outcomes to be considered sound, the measurement of the outcomes must be reliable and valid. A reliable measure

should yield similar results when given to the same person within a short time frame, when they are not receiving an intervention designed to change their performance. A reliable questionnaire should also have items that are internally consistent (i.e., measuring the same characteristic). Both of these forms of reliability reduce measurement error and lead to stronger evidence in support or refutation of a hypothesis. Reducing measurement error also makes it easier to detect a significant effect of a program, thereby reducing the chance that program evaluators will erroneously conclude that a program does not work. Reliability alone, however, does not make a measurement instrument trustworthy. Just because a measure keeps rating a workplace as having a positive climate does not mean that it is a valid instrument. For instance, if a significant number of workplaces rated as having a positive climate by an instrument end up having high turnover rates, unhappy workers, and low productivity, the validity of the scale may be low. The validity of a measure is often established with positive correlations between the instrument and other variables that should be significantly associated with the underlying construct. For instance, an intrinsic motivation (a love for what one does) test may be validated by intrinsic motivation scores being significantly positively correlated with achievement and happiness because both self-determination theory and previous research indicate that intrinsic motivation leads to greater happiness and achievement.

Unfortunately, people routinely use outcomes that may be unreliable or invalid. For instance, many schoolteachers use office referrals (students being required to visit the principal's office) as an outcome to measure the effects of a behavior intervention program. Reduced office referrals are supposedly a sign of better classroom behavior. But some teachers will take pride in the fact that they almost never refer a student to the office, and others will over-refer prior to the intervention but then refer less frequently once the program starts in order to please the principal. This leads to the measures being unreliable and invalid because students' behavior is not the sole determinant of whether they get an office referral or not. It would be better to use empirically validated behavior rating scales with

national norms (clear comparisons to the average boy or girl in the United States) or to use systematic observations by an independent observer (i.e., an observer who is neutral and is not biased toward the students, the program, administrators, or teachers).

Evidence in Support of Theory

In common parlance, the term *theory* is used very lightly, in terms of the evidence most people offer in support of their *theories*. Often, it would be more accurate for people to label their theories as conjecture, assumptions, hunches, or best guesses. Scientists are usually very careful about their use of the term *theory*. In fact, even a hypothesis is not usually formed until one has received years of advanced training in the area of inquiry, rigorously reviewed the scientific literature, and perhaps conducted other related scientific studies under the guidance of well-established scientists. Theories are used to guide the formation of many hypotheses that should be correct if the theory is accurate. Each hypothesis must be tested and replicated in rigorous scientific studies before the theory has adequate evidence. All the aforementioned types of scientific evidence can be used to support the development of a theory, but generally the theory will not be recognized until it has numerous longitudinal and experimental studies to support it. Also, theories must be testable, which makes them capable of being refuted, confirmed, or revised in light of evidence. When one sets forth a theory that cannot possibly be tested, scientists will generally point out that it is not a scientific theory or not a theory at all.

It is important to point out that only scientific studies that are published in peer-reviewed scientific journals are generally considered as evidence or counterevidence related to a theory. Editors of peer-reviewed scientific journals typically send *blind copies* (copies that do not identify the names of the authors) of the manuscripts submitted to three or more established scientists who are either experts in related content or the methods addressed in the study. When the editor chooses to reject a manuscript, it is usually due to major weaknesses in the measures, methods, or review of the scientific literature. This helps to

prevent evidence that is not sound from entering the scientific body of literature.

Many infomercials, product websites, and flyers claim that their product leads to greater vitality, fitness, health, success, and so forth. However, these commercials often rely heavily on testimonials, including paid testimonials. Although testimonials can be interesting and even quite compelling, the fact is that most consumers do not really need to know if the product works well for one or two people that appear in the advertisement; they need to know whether there is evidence that it works for most people and how well it usually works. They may even want to know if it usually works well for people of similar backgrounds and ages. This is where scientific studies help to provide the type of evidence consumers truly want. Some product developers avoid the scrutiny of peer-reviewed journals because they fear that the lack of a significant positive effect would be discovered. In other words, they fear that rigorous scientific studies would demonstrate the inefficacy of their products. Most of the *evidence* that many advertisements provide would not be considered as sound evidence in support of a hypothesis, treatment, program, or theory. Advertisers of some products go even further to avoid evidence by simply showing imaginary people (actors playing characters) having fun, getting dates, or being very successful while using a product. In doing this, they are trying to bypass reasoning about evidence altogether, hoping that consumers will simply identify with the characters and want the rewards that are putatively associated with the characters and the product.

Lack of Evidence Versus Negative Evidence

It is very important to distinguish between a lack of evidence to support a product, program, or hypothesis and the existence of evidence showing that product, program, or hypothesis is ineffective or harmful. For instance, on the basis of a plethora of rigorous scientific studies over the course of decades, the U.S. Surgeon General has confidently declared that smoking cigarettes causes lung cancer, cardiovascular disease, and pulmonary disease. There is also ample research to indicate that Drug Abuse Resistance

Education (DARE) is not an effective prevention program and that there are many other, less popular, drug abuse prevention programs that are effective. In both of these cases, strong forms of evidence indicate that cigarettes and DARE are not worth the investment of financial resources.

In contrast, there are many programs, products, and interventions being used across the world that simply have not been subjected to rigorous scientific study (e.g., longitudinal, quasi-experimental, and experimental studies using reliable and valid measures of outcomes) and then submitted to peer-reviewed scientific journals. Likewise, there are many assumptions, beliefs, and assertions that have not been rigorously tested. It would be misleading to say that those untested beliefs and products are no good or harmful. One can only say that they do not have strong evidence either supporting them or disproving them. In some cases, products and programs have not been tested because the developer is avoiding peer review, but in many cases, they have not been tested because of a lack of understanding of the benefit of testing them and a lack of understanding of the levels of scientific evidence. For example, as mentioned before, many administrators and practitioners do not realize that without employing a control group, there is no way of knowing whether a program or treatment works for most people. There are also some beliefs and assertions that are frankly not testable at this time but may be testable in the future as technology and science advance. It is important to realize that beliefs that are not supported by high levels of scientific evidence at this time may in fact prove to be powerfully correct at some point in the future. For instance, current science is unable to test metaphysical and spiritual claims, with the exception of repeatedly finding that people who are active in their faith are generally less depressed and recover more quickly from diseases, on average. Therefore, it is important to recognize the limits of science in gathering certain types of evidence, when discussing the highest levels and most convincing forms of evidence.

In conclusion, the increased emphasis on evidence in modern society suggests that those who can gather and analyze evidence effectively are likely to become increasingly valuable to corporations,

universities, schools, and communities. However, analyzing evidence well also has personal benefits, such as making wiser decisions about investments in products or programs. Knowing and identifying the types of weak evidence that many people rely on can prevent one from being misguided by unfounded assertions in the realm of politics, advertisements, or proposed community projects. Furthermore, understanding the levels of scientific evidence can help consumers and leaders better evaluate the existing evidence while helping program evaluators in various domains of life to conduct more rigorous evaluations that produce stronger evidence in support or in refutation of hypotheses.

John Mark Froiland

See also Analysis; Assumption; Fact Versus Theory; Inference; Modeling

Further Readings

- Achinstein, P. (2005). *Scientific evidence: Philosophical theories and applications*. The Johns Hopkins University Press.
- Chambless, D., & Ollendick, T. (2001). Empirically supported psychological interventions: Controversies and evidence. *Annual Review of Psychology*, 52(1), 685–716. <https://doi.org/10.1146/annurev.psych.52.1.685>
- Deutschman, C. (2010). *Evidence-based practice of critical care*. W. B. Saunders.
- Devine, F., & Heath, S. (2009). *Doing social science: Evidence and methods in empirical research*. Palgrave Macmillan.
- Froiland, J. M. (2018). Promoting gratitude and positive feelings about learning among young adults. *Journal of Adult Development*, 25(4), 251–258. <https://psycnet.apa.org/doi/10.1007/s10804-018-9294-0>
- Froiland, J. M., & Davison, M. L. (2020). Social perception: Relationships with general intelligence, working memory, processing speed, visual-spatial ability, and verbal comprehension. *Educational Psychology*, 40, 750–766. <https://doi.org/10.1080/01443410.2020.1732873>
- Froiland, J. M., & Oros, E. (2014). Intrinsic motivation, perceived competence and classroom engagement as longitudinal predictors of adolescent reading achievement. *Educational Psychology*, 34, 119–132. <https://psycnet.apa.org/doi/10.1080/01443410.2013.822964>
- Froiland, J. M., & Smith, L. (2012). Advancing the discussion about systematic classroom behavioral observation, a product review of Tenny, J. (2010). eCOVE observation software. Pacific City, OR: eCOVE Software, LLC. *Journal of Attention Disorders*, 18(4). <https://doi.org/10.1177%2F1087054712436585>
- Gettler, L., McDade, T., Feranil, A., & Kuzawa, C. (2011). Longitudinal evidence that fatherhood decreases testosterone in human males. *National Academy of Sciences*, 108(39), 1619–16199. <https://doi.org/10.1073/pnas.1105403108>
- Guyatt, G., Oxman, A. D., Vist, G. E., Kunz, R., Falck-Ytter, Y., Alonso-Coello, P., & Schünemann, H. J. (2008). GRADE: An emerging consensus on rating quality of evidence and strength of recommendations. *British Medical Journal*, 336(7650), 924–926. <https://doi.org/10.1136/bmj.39489.470347.AD>
- Jenkins, S. (2004). *How science works: Evaluating evidence in biology and medicine*. Oxford University Press.
- Kiefer, M., & Pulvermueller, F. (2012). Conceptual representations in mind and brain: Theoretical developments, current evidence, and future directions. *Cortex*, 48(7), 805–825. <https://doi.org/10.1016/j.cortex.2011.04.006>
- Kowalski, M. J., & Froiland, J. M. (2020). Parent perceptions of elementary classroom management systems and their children's motivational and emotional responses. *Social Psychology of Education*, 23(2): 433–448. <https://doi.org/10.1007/s11218-020-09543-5>
- Lijmer, J., Mol, B. W., Heisterkamp, S., Bossel, G. J., Prins, M. H., Van Der Meulen, J. H., & Bossuyt, P. M. (1999). Empirical evidence of design-related bias in studies of diagnostic tests. *JAMA*, 282(11), 1061–1066. <https://doi.org/10.1001/jama.282.11.1061>
- Lilienfeld, S. O., Ammirati, R., & David, M. (2012). Distinguishing science from pseudoscience in school psychology: Science and scientific thinking as safeguards against human error. *Journal of School Psychology*, 50, 1.
- McGee, S. (2012). *Evidence-based physical diagnosis*. W. B. Saunders.
- Melnyk, B. M., & Fineout-Overholt, E. (2010). *Evidence-based practice in nursing & healthcare: A guide to best practice* (2nd ed.). Lippincott Williams & Wilkins.

- National Research Council. (2011). *Reference manual on scientific evidence: Third edition*. The National Academies Press.
- Owen, D. (2000). *Hidden evidence: Forty true crimes and how forensic science helped solve them*. Firefly Books.
- Powell, D. R., Son, S.-H., File, N., & Froiland, J. M. (2012). Changes in parent involvement across the transition from public school prekindergarten to first grade and children's academic outcomes. *The Elementary School Journal*, 113(2). <https://doi.org/10.1086/667726>
- Sober, E. (2008). *Evidence and evolution: The logic behind the science*. Cambridge University Press.
- Strong, W., Malina, R. M., Blimkie, J. R., Daniels, S. R., Dishman, R. K., Gutin, B., Hergenroeder, A. C., Must, A., Nixon, P. A., Pivarnik, J. M., Rowland, T., Trost, S., & Trudeau, F. (2005). Evidence based physical activity for school-age youth. *The Journal of Pediatrics*, 146(6), 732–737. <https://doi.org/10.1016/j.jpeds.2005.01.055>

EVOLUTIONARY PSYCHOLOGY

Evolutionary psychology involves understanding the human mind and behavior from an evolutionary perspective, starting with Charles Darwin's theory of natural selection and including advances in theory and data over the last 160 years. The basic tenet of evolutionary psychology is that the human mind has been prepared by natural selection, operating over geological time, for life in a human group. Evolutionary psychology recognizes that humans are social animals and that much of our evolved psychology concerns how to cooperate and compete with our fellow *Homo sapiens*. Evolutionary psychology as a formal academic discipline is relatively young, dating back only to the last decades of the 20th century. Its roots, however, can be traced back to Darwin's 1859 seminal book, *On the Origins of the Species*, where he wrote, "In the distant future . . . psychology will be based on a new foundation, that of the necessary acquirement of each mental power and capacity by gradation" (p. 488).

Central to any evolutionary theory is the concept of *adaptation*, reliably inherited features that helped organisms survive and reproduce. An

evolutionary perspective on human thought and behavior assumes that not only has the human body been adapted through evolution for survival (e.g., the opposable thumb critical for fine-motor actions necessary in tool use; the digestive system adapted to an omnivore diet) but so has the human mind and behavior. Moreover, adaptations are found across the life span—not only in adults who do the all-critical reproduction but also in infants and children, serving to keep them alive through a perilous time in life and to prepare them for adulthood.

Although an evolutionary perspective can be applied to any areas of human behavior, it is primarily concerned with four general topics: *survival*, including food preferences and predator/prey relations; *reproduction*, including sexual attraction and mate selection; *parenting*; and *social relations*, including learning how to compete and cooperate with the fellow humans, reciprocity, in-group/out-group relations, and status striving. The following sections examine some of the assumptions and principles of evolutionary psychology and then provide examples of human adaptations related to survival, reproduction, parenting, and social relations.

Assumptions and Principles of Evolutionary Psychology

Darwin's Theory of Natural and Sexual Selection

Evolutionary psychology takes as a starting point Charles Darwin's theory of evolution by *natural selection*. Natural selection has four general components. First, within a species, more individuals are born in a generation than will survive. Second, there is variation in form or function among members of a generation. Third, this variation is inheritable. And finally, features that result in individuals surviving and reproducing tend to be selected as a result of the interaction between the individual and the environment. Natural selection is not forward looking or necessarily is the *best* solution to a problem but rather works with what is already available, adapts individuals to their local environment and not to a future one, and is economical so that more individuals with a particular feature will survive and reproduce than

those without it. Darwin also emphasized that selection operated at the individual level, not at the level of the group or species. In the early 20th century, Darwin's theory was integrated with the genetic theory of mathematician, biologist, and Augustinian friar Gregor Mendel (1822–1884), producing what has been called *neo-Darwinism*, or the *modern synthesis*.

In addition to natural selection, Darwin noted that some traits evolved simply because members of one sex found them attractive in the opposite sex, even if they might have been a handicap for the sex that possessed them. Darwin referred to this as *sexual selection*. Consider the peacock's tail. Males with long and beautiful tails are favored by peahens, even though the large tails make the males more susceptible to predation. Yet, large and beautiful tails are honest signs of health ("I must be healthy if I can carry around this tail and still be alive"), a signal to females that such males will likely provide good genes for their babies.

Evolved Psychological Mechanisms

Perhaps the central concept in evolutionary psychology is that of *evolved psychological mechanisms*. These are information-processing mechanisms (i.e., brain-based ways of encoding, processing, and storing information) that underlie thought and behavior and are the *missing link* in evolutionary explication. Early theorists proposed that these mechanisms are *domain specific*, in that they were shaped by natural selection to solve specific recurrent problems faced by our ancestors (e.g., identifying social cheaters; ideal characteristics in a potential mate), analogous to different organ systems in the body (e.g., the lungs evolved for respiration, the heart for pumping blood). These evolved mechanisms should not be viewed as being expressed in exactly the same way in all individuals in all circumstances but rather viewed as predispositions or biases in terms of how certain classes of information should be processed and are then shaped by experience to produce (usually) adaptive outcomes. Thus, evolved psychological mechanisms display a high degree of *plasticity* (the ability to change), permitting individuals to adapt to a wide range of environments. Contemporary theory postulates that humans also evolved some

domain-general information-processing mechanisms, permitting individuals to deal with novelty.

As an example of evolved plasticity, consider recent findings from *life history theory*. Life history theory is a mid-level evolutionary theory concerned with decisions animals make in allocating time and energy to various aspects of their development. With respect to humans, children growing up in environments that are harsh (low income, lack of resources) and unpredictable (frequent change of residence, unreliable and changing social relations) modify their development in anticipation of having similar environments as adults, developing a *fast life history strategy*. They thus take an opportunistic perspective, reaching puberty earlier, engaging in more risky behavior, and becoming sexually active earlier with more partners in comparison with those who, as children, grew up in less harsh and predictable environments, and who generally take a more futuristic approach to life, developing a *slow life history strategy*. Although the life course of children following a fast life history strategy may be seen as maladaptive from a broader societal perspective, it is well adapted to the difficult life circumstances such children can anticipate in adulthood.

Environment of Evolutionary Adaptedness

Evolutionary psychologists refer to the time and ecology when an adaptation arose or became established as the *environment of evolutionary adaptedness*. Roughly speaking, it refers to the environment to which a species is adapted. Humans are mammals and thus share a long evolutionary history with all other mammals, along with adaptations. Humans last shared a common ancestor with chimpanzees and bonobos between 5 and 7 million years ago (mya), with fossil evidence of the first member of the human family (*Homo habilis*) dating back about 2.5 mya, with anatomically modern *Homo sapiens* appearing in Africa as early as 300,000 years ago. With respect to unique human psychological adaptations, most evolutionary psychologists focus on the last 300,000 years, when our ancestors made their living as nomadic hunter-gatherers on the savannas of Africa. Fossil and archeological evidence, plus information about how modern hunter-gatherers

live, give us an idea of what the life of our ancestors might have been like.

Our ancestors likely lived in small, nomadic groups of between 30 and 100 people, interacting occasionally with other groups, trading mates and goods. They obtained nutrition mainly by gathering (mostly the work of women), hunting (mostly the work of men), and scavenging. Females likely reached puberty and had their first births in their late teens or early 20s, nursed infants for about 3 years, and gave birth every 4 to 5 years. Child-care was done primarily by mothers with the help of other women in the group, with fathers providing protection and nutrition for their mates and children. Early humans were probably marginally monogamous/polygamous, with some men having several *wives*, although most men likely had a single wife, with some men having no access to women. The physical environment of Africa, and later of Europe and Asia, was highly variable over the past 300,000 years, meaning ancient humans had to be able to adapt to substantial changes in climate and habitat. *Homo sapiens* developed agriculture and began to live a sedentary lifestyle about 10,000 years ago. Since that time, the pace of cultural evolution has outstripped that of biological evolution, and because of this, many adaptations that evolved in ancient environments and are inherited by contemporary humans are sometimes mismatched to modern environments. For example, humans' preference for sweet and fatty foods made good sense when people did not know where their next meal was coming from but may be maladaptive for people in industrialized societies living with fast-food restaurants and grocery stores. In addition, just because an adaptation evolved in the environment of evolutionary adaptiveness or is based in our biology does not make it *good*, *right*, or *justified*, nor does it mean that it is inevitable. This is the *naturalistic fallacy*, and human behavior evolved to be sufficiently plastic (changeable) so that biology need not be destiny.

Survival

Knowing What to Eat and What Not to Eat

Humans are omnivores, eating fruits, vegetables, and meat, and like all omnivores must learn what is okay to eat and what potential foods must

be avoided. Humans' evolved preference for sweet and fatty foods was mentioned earlier. Such taste cues are signals of foods high in calories and thus nutrition. Other tastes and odors, being cues of potentially toxic substances, seem to result in avoidance, with pregnant women being especially sensitive to certain tastes and smells. For example, many women in the early stages of pregnancy develop aversions to certain types of foods, including bitter-tasting vegetables (for instance, broccoli), meat (especially raw or undercooked meat), and alcohol, all substances that, particularly in ancient environments, could interfere with the development of the fetus they are carrying. Infants acquire some early taste preferences based on the diets of their mothers, both while still in utero through the placenta and after birth through breast milk. For example, in research where mothers ate anise-flavored food several weeks before giving birth, their babies showed a preference for anise-flavored milk after birth, relative to nonexposed infants. Humans' potent social learning abilities also result in children learning to eat what their parents and other adults in the group eat, making it likely they will develop food preferences similar to those of their parents.

Nonetheless, even when adopting the same eating habits as more experienced others, food sometimes makes us sick. Given that there is often an extended delay between consuming food and experiencing sickness, how do people know what it was that made them sick and avoid it in the future? Humans and other omnivores, such as rats, seem prepared to associate sickness with novel foods they have eaten. Such *prepared learning* is illustrated in experiments with rats in which they drank artificially sweetened water (a novelty for them) and were subsequently exposed to low levels of radiation that made them sick hours after consuming the water. After recovering, the rats avoided the novel water when given an option. In this respect, humans are no different, and this has sometimes been referred to the *sauce béarnaise effect*, reflecting the belated nausea that spoiled béarnaise sauce can have and people's subsequent aversion to it.

Developing Fear of Predators

For our ancestors, staying alive meant avoiding being eaten by carnivores or otherwise killed by

dangerous animals. Snakes remain the vertebrate causing most human death in the world today, and they were surely a threat to our ancestors. Fear of snakes seems to be universal, with many individual differences, although such fear seems not to be innate but rather easily acquired when infants and young children watch someone respond fearfully to snakes. For example, in one research project, infants and toddlers viewed short videos of snakes and other animals while they listened either to a pleasant or a fearful voice. The children looked significantly longer at the snakes when they were paired with a fearful voice than with a pleasant one. There were no differences in their attention when pleasant and fearful voices were paired with other potentially dangerous animals such as an elephant or hippopotamus. Apparently, snakes' sinusoidal movement, different from the movement of mammals, attracted infants' attention and became an associative cue in children's developing a fear of this prehistorically dangerous class of animal. Other research has similarly shown that monkeys have no innate fear of snakes, but, like human children, can easily acquire such fear by watching another monkey respond fearfully to snakes. They do not develop fear responses when they watch another monkey respond fearfully to a rabbit or a flower.

Reproduction

Parental Investment Theory

The essence of evolutionary theory is reproduction, passing one's genes to the next generation. In humans and most animals, reproduction is done sexually. Although at one level males and females have the same goal (having offspring that grow up to be reproductive adults themselves), the self-interests of the two sexes are not identical. This is best understood from the perspective of *parental investment theory* as initially articulated by Robert Trivers in 1972.

The principal concept of parental investment theory is that there is a conflict for both males and females in how much time, effort, and resources to invest in mating versus parenting. In most mammals, including humans, there is greater obligatory investment in parenting for females than for males. Greater female investment begins with the

size of gametes (eggs being substantially larger than sperm), followed by internal fertilization and gestation, and nursing, which, until the modern era, was the only source of nutrition for young infants. Females in all cultures also spend more time in childcare than males do. According to the theory, this differential investment in parenting between the sexes results in the sex that invests less in parenting (usually males) being more willing to engage in sex and being more competitive with members of the same sex for access to members of the opposite sex. This pattern generally reflects sex differences in behavior in most mammals, including humans.

For example, women are more hesitant to consent to sex and are less desirous of having multiple sex partners than men are. A single sex act can result in 9 months of pregnancy for a woman and several more years when she is the sole source of nutrition for her baby. A man's potential contribution can be as little as his sperm and the time it takes to copulate. Many men do commit to women and their children, however, and they are most apt to do so when such commitment is necessary for the survival or social success of their children, and when it is necessary to attract a high-quality mate (commitment is what women want).

The higher rate of aggression, including homicide, for adolescent and young adult males than for females reflects males' greater competition for social status, which women value in men. Although most women will find a mate, even if a less-than-desirable one, some men in some societies can be shut out of mating entirely, making it worth their while to engage in reckless and often potentially deadly actions.

Sex Differences in the Ideal Mate

In most cultures, men and women want the same thing in a mate: someone who is kind and loving, healthy, and compatible in interests and personality. But there are typical cross-national, pan-cultural differences in the features that males and females desire in an ideal mate, and these can be explained (and predicted) by evolutionary theory. For example, men more than women value attractiveness and youth in a potential mate. The reason for this seems to be that both

attractiveness and youth are associated with health and a woman's *reproductive value*—how many children she can have in the future. For instance, degree of facial symmetry (the extent to which the left and right sides of the face are similar) is a major contributor to people's ratings of attractiveness and a reflection of low levels of perturbed development and thus health. Moreover, men in most societies find women with an hourglass shape (a waist-to-hip ratio of about 0.7) the most attractive, and this shape has been found to be related to fertility in cultures that do not use hormonal contraceptives. In contrast, women more than men in most cultures value high social status and access to resources, such as food and money, as more important in a potential mate. This reflects the ability of a man to invest in her and her offspring.

Jealousy

Infidelity is a problem in any relationship, and becoming jealous of a potential competitor for one's partner's attention can be adaptive, in that it can promote behavior aimed to enhance the relationship (being more attentive to one's partner) or behavior aimed to punish via aggression toward the partner or the potential competitor.

Jealousy comes in (at least) two forms—sexual and emotional—and men and women tend to respond differently to the two types. Sexual jealousy involves concerns that one's partner is having sex with someone else, whereas emotional jealousy involves concerns that one's partner is becoming emotionally involved (without having sex) with someone else. Both men and women are disturbed by both types of jealousy, but both verbal reports and physiological measures (blood pressure, galvanic skin response) show that men are more bothered by a partner's sexual infidelity and women more bothered by a partner's emotional infidelity. Men's greater distress with a partner's possible sexual infidelity is explained by *paternity uncertainty*. Prior to the advent of genetic testing, a man could never be 100% certain that he is the genetic father of a child born to his partner, and if his partner were having sex with another man, this can result in *cuckoldry*, a domestic father raising a genetically unrelated

child. In contrast, a woman always knows that a baby is hers, so the consequences of sexual infidelity are not as great for a woman. (She's not about to unknowingly raise another woman's child.) However, if her partner becomes emotionally involved with another woman, this could lead to him abandoning her and her children for another woman, and this, traditionally, has been a greater risk for women.

Parenting

Investing in Infants

Infant and child mortality rates for our ancestors were likely about 50%. Ancestral women thus had to carefully evaluate the likelihood of any child surviving to adulthood and invest in each child accordingly. Women who devoted too many resources to a sickly infant, for example, jeopardized the health and survival of any current children she had as well as any future children she might have. There are a variety of factors that ancestral (and many modern) women must evaluate to determine how much (or how little) to invest in an infant. Perhaps most important is a child's health. In traditional cultures, sickly infants are apt to receive less investment and are more apt to be abandoned than healthier infants. For instance, if twins are born and the mother lacks the resources to raise both, the weaker of the two is likely to be abandoned. Younger children are more likely to be abandoned or given fewer resources than older children because the older children have already survived to a greater age and thus are more likely to reach adulthood than younger children. Younger women are more likely to abandon an infant than older women because they will have more subsequent opportunities to become pregnant and have children, and women with greater social support (a mate, family) are more likely to invest heavily in an infant than women with less social support. Although we may like to believe that women (and men) love all their children equally and will invest whatever is needed for their survival, the truth for ancestral women (and to a lesser extent men) is that they had to be discriminating in how much to invest in any child, and it was such discriminating women who became our foremothers.

Adaptations of Infants to Garner Adult Attention

Infants, however, are not passive recipients of their parents' attention but have evolved a set of psychological *weapons* to obtain attention from adults to increase their chance of survival. For instance, infants are attentive to biological motion, are especially attentive to faces, and are responsive to adults' facial expressions, all of which encourage adults' caregiving. Infants' large heads, eyes, and other features of babyhood promote positive feelings and caregiving in adults, particularly in women and especially for a woman's own infant, again increasing the likelihood of mother–infant attachment and thus infant survival. Recent research has also shown that infants as young as 9 months of age will show *jealousy protest* when their mothers pay attention to another infant in contrast to when they attend to an inanimate object such as a book. Another infant is a potential social rival, and jealousy protest may serve to decrease the chance of an infant being displaced by a competitor.

Social Relations

Humans Are a Hypersocial Species

Humans are a highly social species but so are chimpanzees, wildebeests, wolves, and baboons, among many others. But humans' sociality is extreme—we are *hypersocial*, displaying a remarkable ability to cooperate with others to achieve tasks that a single (or even a handful) of people could not accomplish. Human cooperation extends beyond family and friends, often involving large groups of people, some of whom might not be known well or at all. Human sociality involves notions of morality, often helping others and treating others fairly at some expense to oneself, which is found in all cultures (although the specific moral prescriptions can differ substantially among societies). Darwin had difficulty explaining humans' sense of morality and seeming acts of altruism and suggested that human morality may be an exception to his rule that natural selection only operates on the level of the individual and indeed may, in this case, operate on the level of the group.

The nature of human altruism and *Homo sapiens'* hypersociality has been a contentious one ever since Darwin. One current account, *multi-level selection theory*, holds that natural selection operates at a variety of levels, the group being one of them. Specifically, groups of individuals showing high levels of cooperation function more efficiently than less cooperative groups and are more apt to survive and thrive. From this perspective, selfish individuals are likely to outcompete altruistic individuals within a group, but altruistic and cooperative groups are likely to outcompete selfish groups.

Humans' tendencies toward cooperation can be seen early in development. Although 3-year-old children are not known for their tendency to share and treat others fairly, when they work together to obtain some reward, they tend to distribute the spoils of their efforts evenly. Highly social chimpanzees do not do this (nor do human 2-year-olds), and as a result, collaborative efforts between chimpanzees quickly decline when one member of a pair consistently takes the larger portion of the rewards.

This is not to say that humans always cooperate with one another and are always considerate of others. Both selfishness and cooperation characterize *Homo sapiens*, and the social adaptations necessary for altruism and high levels of collaboration evolved from evolutionary adaptations that humans share with other primates. Along these lines, in order to understand human sociality, it is necessary to contrast it with that of our close genetic relatives, chimpanzees and bonobos, who are the best models we have for what our last common ancestor with these animals may have been like. Extending this argument, Michael Tomasello has further proposed that it is necessary to focus on how the development of great ape sociality has been transformed into the development of human sociality.

Adapted to Social Life

Much like other social animals, humans have a strong tendency to affiliate with other group members and as a result tend to show favoritism toward in-group members and fear of/or hostility toward out-group members. Experiments have shown that very little is needed for people to identify with a group and show favoritism toward that group.

Simply being assigned different colored T-shirts can result in group identification, with in-group members (people wearing the same color T-shirts) being viewed more positively and receiving more benefits than out-group members (people wearing different color T-shirts). Moreover, such in-group/out-group distinctions based on this *minimal groups paradigm* is observed even among preschool children. People, again beginning in early childhood, tend to pay attention to and adopt social norms, often penalizing people who break such norms.

Humans, like other social animals, are also status seekers. Individuals who have higher standing in the group often have access to more resources, which can translate into enhanced mating opportunities and benefits to themselves and their offspring.

Nothing in Psychology Makes Sense Except in the Light of Evolution

An evolutionary perspective provides a meaningful context in which to make sense of human cognition and behavior. An evolutionary perspective does not deny the role of experience or culture in shaping human behavior. Humans are, and always have been, a cultural species, and they have evolved adaptations for living in cultural groups. Nor does an evolutionary perspective imply *genetic determinism*, the idea that genes determine behavior. The ability to adapt to changing environments (plasticity) is an evolved feature of *Homo sapiens*. An evolutionary perspective on human psychology serves to organize known facts economically; provides guidance to important domains of human functioning, including those associated with modern culture such as psychopathology and formal education; leads to new predictions about human behavior; unifies psychology with the other life sciences (biology, medicine); and can serve as a *metatheory*—a common set of broad, overarching assumptions and principles—for other areas of psychology.

David F. Bjorklund

See also Biology, Evolutionary; Natural Selection

Further Readings

- Barkow, J. H., Cosmides, L., & Tooby, J. (Eds.). (1992). *The adapted mind: Evolutionary psychology and the generation of culture*. Oxford University Press.
- Bjorklund, D. F. (2021). *How children invented humanity: The role of development in human evolution*. Oxford University Press.
- Buss, D. M. (2011). *The dangerous passion: Why jealousy is as necessary as love and sex* (2nd ed.). The Free Press.
- Buss, D. M. (2016). *The evolution of desire: Strategies of human mating* (4th ed.). Basic Books.
- Buss, D. M., Abbott, M., Angleitner, A., Asherian, A., Biaggio, A., Blanco-Villasenor, A., Bruchon-Schweitzer, M., Ch'U, H.-Y., Czapinski, J., Deraad, B., Ekehammar, B., Lohamy, N. E., Fioravanti, M., Georgas, J., Gjerde, P., Guttman, R., Hazan, F., Iwawaki, S., Janakiramaiah, N., . . . Yang, K. S. (1990). International preferences in selecting mates: A study of 37 cultures. *Journal of Cross-Cultural Psychology*, 21, 5–47.
- Daly, M., & Wilson, M. (1988). *Homicide*. Aldine.
- Darwin, C. (1859). *On the origin of species*. Modern Library.
- DeLoache, J. S., & LoBue, V. (2009). The narrow fellow in the grass: Human infants associate snakes and fear. *Developmental Science*, 12, 201–207. <https://psycnet.apa.org/doi/10.1111/j.1467-7687.2008.00753.x>
- Dunbar, R. I. M. (2010). *How many friends does one person need? Dunbar's number and other evolutionary quirks*. Faber & Faber.
- Ellis, B. J., Figueredo, A. J., Brumbach, B. H., & Schlomer, G. L. (2009). Fundamental dimensions of environmental risk: The impact of harsh versus unpredictable environments on the evolution and development of life history strategies. *Human Nature*, 20, 204–268. <https://doi.org/10.1007/s12110-009-9063-7>
- Garcia, J., & Koelling, R. A. (1966). Relation of cue to consequence in avoidance learning. *Psychonomic Science*, 4, 123–124.
- Geary, D. C. (2021). *Male, female: The evolution of human sex differences* (3rd ed.). American Psychological Association.
- Tomasello, M. (2019). *Becoming human: A theory of ontogeny*. Belknap Press.
- Trivers, R. (1972). Parental investment and sexual selection. In B. Campbell (Ed.), *Sexual selection and the descent of man* (pp. 136–179). Aldine de Gruyter.
- Wilson, D. S., & Wilson, E. O. (2007). Rethinking the theoretical foundation of sociobiology. *Quarterly Review of Biology*, 82, 327–348. <https://doi.org/10.1086/522809>

EXPERIMENT, THEORY OF

Experiments to understand and improve agriculture and animal husbandry must be as old as the history of humankind once we had ceased to be exclusively hunter-gatherers. Yet even the great advances in these areas in the late 18th century were achieved without any general theory of experiments. The same holds for the experiments of Renaissance scientists such as Galileo, of many 18th- and 19th-century chemists and of 19th-century experiments in science and engineering. These often took place in laboratories under carefully controlled conditions; typically, only one factor was varied at a time. Toward the end of the 19th century, in North America, Europe and its colonies and dependencies, and Japan, there was a marked development of research stations devoted to agricultural experimentation. However, it was not until the second decade of the 20th century that a theory emerged, leading to efficient experiments with many factors that allowed for the haphazard fluctuations that come with the measurement of natural systems. The first systematic understanding and discussion of this theory seems largely to have been due to Sir Ronald Fisher, indeed working at an agricultural research station. His contributions to the theory of design of experiments and their analysis are the subject of the next section. Thereafter, this entry traces the development of these ideas as the theory extended to cover technological and computer experiments. In what follows, the topics are introduced in historical order.

Sir Ronald Fisher and Agricultural Experiments

The Rothamsted Experimental Station was started in 1843 by John Bennet Lawes on his estate, Rothamsted Manor, in Hertfordshire, England. Lawes had founded one of the first artificial fertilizer manufactories. His initial purpose was to investigate the impact of inorganic and organic fertilizers on crop yield, at a time when it was still common belief that organic chemicals differed in the presence of a *vital force* from those that were inorganic (the synthesis of urea dates from 1828). Lawes launched a series of long-term field experiments, some of which continue to this day. Over

the next half-century, Lawes and his scientific adviser Joseph Gilbert established the foundations of modern scientific agriculture and the principles of crop nutrition.

In 1919, the young Ronald Fisher (he was born in 1890) was employed at Rothamsted to investigate the possibility of analyzing the data resulting from the station's crop and grasslands experiments. This experience led him to the first systematic formulation of a theory of experimental design. Because of this background, much of the nomenclature for the subject comes from that of the agricultural field trial.

The central problem in the theory of the design of experiments is which treatments to apply to which units. The units may be plots in a field, patients, batches of raw material, and so forth. The treatments are procedures under the control of the experimenter that are to be applied to individual units. In agricultural experiments, they might be the amounts per square meter of nitrogen, phosphorus, and potassium (NPK), perhaps together with numbers and times of application. In a clinical trial, the treatment is often the dose of a particular drug, administered in a particular way, such as by tablet or injection. In cereal experiments, the response is usually the yield of grain, although secondary responses, such as the length of wheat stalks, may be important in identifying varieties that will not be damaged by storms, a form of robustness discussed later under the industrial contributions of Genichi Taguchi. In a clinical trial, the response might be reduction in blood pressure, or in number of malarial parasites, following treatment.

Fisher's contributions can be summarized under four headings.

1. *Randomization and Systematic Error.* The treatments need to be allocated to the units in such a way as to avoid biases. If a particular treatment is allocated to an especially favorable set of units, the comparisons with other treatments will be biased. Widespread empirical evidence shows that haphazard allocation by experimenters can lead to such biases and is almost always unsatisfactory. An objective randomization procedure is necessary for the establishment of results that can be confidently extended to other, seemingly similar, situations. This is particularly true, for example, in clinical trials, where the

effects to be measured may be small and subjective reactions of patients and assessors may obtrude.

2. *Control of variability.* If there is appreciable variability between units, accuracy of comparisons between treatments will be improved if they can be made within sets of units (*blocks*) with similar properties. The analysis seeks to estimate the differences in treatment effect accounting for variation between blocks, making sharper comparisons possible. The simplest case is when all treatments are allocated to each block. In agricultural field trials, the blocks will be neighboring parts of a field, although the shape and extent of the block may be determined by the geology and agricultural history of the field. The allocations of treatments within blocks should be randomized, partly to avoid any systematic bias due to unexpected inhomogeneity between plots within blocks.

3. *Independent replication.* The effect of treatments has to be assessed against the intrinsic variation in the response to the same treatment applied under identical conditions to similar units. This can be achieved by repeating some treatments within blocks and, at the cost of further assumptions, by the use of models in analyzing the data. An important assumption is that of unit-treatment additivity, stressed by Jerzy Neyman. Of course, modeling assumptions can and should be checked during the analysis of data.

4. *Factorial designs.* The efficiency of an experiment can be greatly improved by asking several questions at once. The most powerful example is the factorial experiment. To continue with the example of field trials, suppose that interest is in the effects of levels of NPK, which are all to be applied at a high and a low level. The exact definition of these levels will depend on previous experiments and background knowledge. In the experiment, there are eight distinct combinations of treatment levels. When these eight levels are all repeated the same number of times, the resulting design is *orthogonal*; under simple assumptions, it is possible to estimate the effect of each treatment as if the others were not there; estimates of all three treatment effects are obtained which are just as precise as if each of the other two factors had been held constant—a *one-at-a-time* design.

The second powerful feature of the factorial experiment is that it allows detection of interactions which arise if the effect of one factor depends on the level of one or more of the remaining factors. Such important effects occur frequently in experimental work. Crucially, for the theory of design, they cannot be detected from experiments in which only one factor is varied at a time.

Fisher's main contribution to the analysis of experiments is the collection of techniques known as the *analysis of variance*. This was developed at the same time as his theoretical and practical work on experimental design. It comes from the use of least squares to fit a series of models of increasing complexity to data. Although the derivation of least squares goes back to Carl Friedrich Gauss and Adrien Marie Legendre in the first decade of the 19th century, Fisher emphasized the division of the total sum of squares of the observations into components corresponding to estimated treatment effects, to differences between blocks, and to random error. The assumption of normally distributed random errors leads to the distributions of the statistics for testing the values of the estimated treatment effects. For orthogonal designs, he provided straightforward calculation formulae, which remain widely taught and used, and geometrical interpretations in terms of Pythagoras's theorem. He also argued that interpretation of the calculations was little affected by small departures from the assumption of normality.

This approach is based on modeling. Another approach, which does not require a model, reuses the existing data from the randomized experiment. Given a statistic calculated from the data, its significance can be obtained by repeatedly re-randomizing the observed responses to the design and calculating the value of the statistic for each randomization. The significance of the observed statistic is its position in the ordered values of the statistics generated by this series of re-randomizations. Although the approach is model free, the power of the randomization test will depend on the choice of statistic. The method depends on the validity of the assumption of additivity of the treatment effect to that of the unit, where the unit effect is the baseline that could be expected in the absence of any treatment. If this assumption does not hold, it may often be achieved by data transformation.

For example, a logarithmic transformation renders a multiplicative effect additive.

Theoretical results, such as the Neyman–Pearson lemma, indicate the statistic for the most powerful test, under the conditions of a specific model. Quick graphical checks of models are rendered easier to interpret by the balanced treatment allocations of designed experiments. For example, in a factorial experiment, plotting the responses, or their means, against the levels of each of the factors in turn often indicates whether an interaction is present; in the absence of interaction the results for the differing treatments differ by the same amount, within experimental error, whatever the level of the plotting factor. Changes in variance with the level of the factor are an indication of heteroscedasticity, which can often be removed by transformation of the data. These suggestions for model enhancement are then followed by formal model fitting, parameter estimation, and hypothesis testing.

Perhaps the most important aspect of Fisher's work on the design of experiments is the emphasis on randomization to avoid sources of potential bias. But well-designed experiments, following his precepts, are frequently also easy to analyze. As a consequence of balance and orthogonality, a few simple graphics often reveal the important aspects of the effects of treatments. Haphazard experiments, apart from their inefficiency and susceptibility to bias, may lead to poorly identified treatment effects. Even if an effect is evident, the lack of orthogonality may make it impossible to ascribe the effect to changes in specific factors.

Clinical Trials

Clinical trials are experiments to determine treatments for patients. Although they are very different from agricultural trials, the theory described earlier makes a vital contribution, but with a rather different emphasis. In Phase III of such trials, a few treatments are compared. Often just two, a standard treatment and a new one that may offer medical advantages or may be not worse than the standard but cheaper, or preferable in some other way, for example, the absence of side effects. In such experiments, the effects are likely to be small compared with the variation in response

in the population of patients. Random allocation of treatments to patients is important, perhaps after patients have been stratified, for example, by gender or initial disease severity. The randomization needs to be blinded to the physician determining treatment, as well as to the patient. In addition, it needs to be blinded to medical personnel who are involved in the care of the patients and the assessment of their subjective feelings of well-being. Such *double blinding* is widely considered necessary by regulatory agencies to help ensure the objectivity of the results of the trial.

Patients can be recruited for a trial and all be randomized to treatment before the trial starts. Or they can be recruited sequentially and again randomized to treatment. Particularly in sequential trials, the randomization needs to achieve some balance over the prognostic factors with which the patients present. Randomization is usually over a few factors. It may turn out, after the experiment, that there are some imbalances in allocation, for which a specific regression method, the analysis of covariance, can be used to make adjustment.

Industrial Experiments: George Box and the Response Surface

One branch of the theory of design for industrial experiments was developed, after the Second World War, by George Box and colleagues, initially at Imperial Chemical Industries in Manchester, England. In the kind of experiments they encountered, the response was a smoothly varying function of continuous factors, such as temperature, time, chemical concentration, or stirring speed. In some experiments, the aim was to find the maximum over the factor values of such responses as the yield of product. In others, a model for the response was required over a known region: Economic modeling would then establish the best conditions of operation as prices of energy and the input chemicals varied. There is less emphasis on qualitative factors than in agricultural experiments. Furthermore, the units are often more uniform so that replication is of lesser importance in industrial experiments.

Typically, second-order polynomials provide satisfactory models. To fit such models the experiment needs to provide observations at, at least,

three levels of each factor. However, three-level factorials with each factor at a low, intermediate, and high level may rapidly become too large. For three factors, the design requires observations to be taken at 27 factor combinations, while the full model contains only 10 parameters (a constant, the linear and quadratic effects of each factor, and the three two-factor interactions). For four factors, there are 81 factor combinations to estimate a model with 15 parameters. Such designs, because of their large size, may waste resources. The parameter estimates are often more precise than are needed for determining important effects; part of the resources might better be spent on other experimental programs.

Box's contribution to the theory was to focus on the purpose of the design, in his case estimation of the response, rather than on such properties of the design itself as orthogonality. This led to the idea of rotatability, in which designs were found for which the variance of the estimated response increased away from the center of the design in a way that did not depend on the direction. The variance was thus spherically symmetrical. Designs providing such a variance function for a second-order model included two-level factorials augmented by observations taken at the center of the design region and *star points*; these are design conditions in which the experimental settings for all except one factor are held at the central value. The scaling of the design region and the number of center points are determined by the importance of the variance of the response relative to concerns about bias from an inadequate model.

These designs are much smaller than three-level factorials. For three factors, the response surface design contains 17 trials, if three center points are included, compared with 27 for the factorial design. For four factors, the numbers are 27 and 81, with the ratio continuing to decrease as the number of factors increases. The designs are widely used in the process industries, together with two-level factorials and their fractions. The fractional factorial designs allow estimation of main effects and low-order interactions. One application is to provide relatively small response surface designs as the number of factors increases, through replacement of the full two-level factorial by a fractional design.

A disadvantage of such designs in some circumstances is that they are *exact* designs; they give a

specific number of observations at combinatorial combinations of particular factor levels. As such they lack the flexibility of the *approximate* designs, described in the next section, which can be developed from the theory of Jack Kiefer. One example of the comparison of the designs resulting from the two theories arises in blocking response surface designs. Although blocking is usually less important in industrial experiments than in agriculture, batches of raw material, shifts of operators, and days of operation are often sources of heterogeneity that should be guarded against by blocking. In such designs, the sizes of the blocks are typically constrained by the number of units available within each block. The *approximate* theory can provide a design respecting these constraints which is also efficient for estimation of the response surface. However, the division of the set of experimental conditions for a rotatable design between blocks, with the requirement of orthogonality between blocks, can lead to designs which may not respect the number of units available for each block. The required design structure is then determining the scale and nature of experimentation, which should instead reflect economic and scientific goals.

A second contribution of Box to the theory of experimental design was the introduction of evolutionary operation (EVOP) in which industrial plants or processes are run not only to produce product but also to produce information. The basic insight was an analogy with the evolution of species in which small changes are continually occurring; only those changes that are beneficial are retained. The transfer to experiments is to operate at a series of conditions that are small variations on what are believed to be the best operating conditions. If these change slightly, the series of EVOP experiments will indicate the change to the new conditions of best operation. The variations in operating conditions need to be sufficiently slight that quality of product is maintained. However, as in natural evolution, the number of experiments is large so that even small effects will eventually be detected.

Jack Kiefer and Optimum Experimental Design

At much the same time as Box's work in the chemical industry, Jack Kiefer at Cornell University

approached experimental design from a more abstract point of view, which led to the powerful general theory of optimum experimental design. Instead of focusing on just one property of a design, Kiefer considered functions of the information matrix. The inverse of this matrix provides the variances and covariances of the parameter estimates; an experimental design maximizing the determinant of the information matrix minimizes the volume of the normal theory confidence region for the parameter estimates. Other functions of the information matrix lead to designs minimizing the maximum or average variance of prediction over a specified region, to designs focusing on good estimation of a few parameters, or to designs for choosing between models.

A major step forward in the theory was the use of *approximate* or *continuous* designs in which experimental effort was represented as a probability distribution over the points of the design. Combined with the convex nature of the design criteria, these continuous designs satisfy a *General Equivalence Theorem*. This provides easily verified conditions for the optimality of proposed designs and to efficient algorithms for the computation of many kinds of optimum design. The discrete designs needed for applications are found by integer approximation to the continuous design. The efficiency of these approximations is calculated relative to the optimum continuous designs and the most efficient approximation chosen.

This theoretical framework has proved extremely powerful in the construction of efficient designs in a wide variety of contexts. Blocking, mentioned earlier as being problematic for rotatable designs, is achieved through addition of extra parameters to the model describing the data. Numerical algorithms can be used to divide the trials between the blocks in the most efficient way, taking account of any constraints on the maximum size of the individual blocks due to the availability of material. In this way, the design is determined by the available resources rather than the design determining the resources that are required.

Many of the standard designs, such as two-level factorials, are optimum, often according to more than one criterion. These designs require numbers of trials that are powers of two. However, designs for two-level factors may be required

for other total numbers of trials. These can be found numerically; despite a slight loss in orthogonality, these designs are often of surprisingly high efficiency. Likewise, optimum designs can be found for irregular design regions and for nonstandard regression models, such as those in which the values of the factors are constrained by being the components of a mixture. The most appropriate numerical algorithm will depend on the particular problem. Numerical optimization methods for continuous functions are often most appropriate for response surface designs, whereas combinatorial methods are effective for the blocked treatment designs encountered in agriculture.

In linear regression models, such as those used in response surface studies, the parameters enter linearly. The theory can also be extended to finding optimum designs for more complicated models, including the models with nonlinear parameterization that result from chemical kinetics or from some dynamical systems. Further extensions include designs for generalized linear models, in which the response is not normally distributed; binomial and Poisson responses are of particular importance.

An interesting theoretical aspect is that optimum designs for nonlinear models depend upon the values of the parameters the experiment is intended to estimate. This must come from some form of prior information. Sequential experiments, in which the experiment is guided by the information accruing about the true value of the parameters, provide an objective source of prior information. Subjective prior information can come from the collective experience of scientists. In either case, the information can be used in one of two ways. In truly Bayesian designs, the maximization of the design criterion allows for a Bayesian analysis of the experimental results. In the more often used *pseudo-Bayesian* approach, the design criterion is simply maximized over the prior distribution of the parameters, ignoring the knowledge about the distribution of the responses given by the prior distribution. Much current work on experimental designs in medicine, genomics, and engineering is for nonlinear models. The elicitation and methods of incorporation of prior information are an important aspect of the design of efficient experiments in these fields.

Robust Designs, Computer Experiments

The theory outlined in the preceding section is for experiments that lead to understanding the effect of treatments or to modeling physical phenomena. In this context, a robust experiment is one that yields informative answers even if the physical situation is not quite what was expected. An example is the pseudo-Bayesian design of an experiment to estimate the parameters in a nonlinear model. The design is intended to have reasonable efficiency over a range of values of the parameters.

A second use of the term *robust design* is associated with the work of the Japanese engineer Genichi Taguchi, whose experiments aim at determining the robustness of products and processes to random variation through the inclusion of *noise factors* in experiments. One set of such investigations involves conditions of use: A car engine needs to work well not only in the carefully controlled laboratory conditions under which it is developed but also in a variety of weather conditions from very cold to very hot, from dry to humid, and from sea level up to heights at which oxygen levels are reduced. Such factors should be included in the experimental investigation of the properties of the engine. Similarly, machines may not be properly maintained and need to be insensitive to imprecise adjustments.

An extension of these ideas is to the manufacturing process itself. Here the idea, sometimes called *robust engineering design*, is to design the product so that its performance is insensitive to slight variations in the manufacturing conditions, such as might arise from batches of raw material, or production in different plants. In the analysis of possible designs, there are typically a large number of factors which may interact; the theory of multifactor experiments is crucial to finding a path to the product with greatest engineering robustness. The response to these experiments is typically the proportion of products that do not perform within specification. The analysis, and so the design, therefore involves generalized linear models rather than the normal theory models underlying Taguchi's statistical recommendations.

In robust design for products, rather than in the manufacturing process, the basic statistical

ideas of design are the same as those already described here. Improvements can be made in many of the original proposals of Taguchi by using more recent developments in experimental design, such as those associated with the methods of optimum design. This is in contrast with computer experiments which are methods for exploring the output of a deterministic system and so require a different theory from which to develop informative experiments. Many deterministic computer models, used for example in engineering design, involve large sets of differential equations with complicated boundary conditions. Almost invariably there are many inputs. There may also be many outputs, but this is less critical for the design of efficient experiments. A single run of the simulation model under one set of inputs can often take several hours. Because the model is deterministic, replicate experiments at the same combination of factors yield the same response. However, the outputs usually vary in a smooth way with the inputs, with adjacent inputs having similar output values. It is thus possible to build a simple model, an emulator using a derived response surface, from which the outputs can be determined for any specified set of inputs. Because there is no error of observation, the standard methods of optimum design, such as D-optimality, with their resultant emphasis on replication, are not applicable. One approach to experimental design in these conditions is to use designs that fill the space of the factors as precisely as possible. The theory underlying the designs is close to that used in determining good sequences of pseudo-random numbers that are, in fact, often required in stochastic simulations. A second approach is to assume that the responses form a Gaussian random field. Prior specification of the parameters of the field is required. A Bayesian design is then found which provides best average prediction over the region of interest.

This theory is for simulations without any stochastic component. Simulations with a stochastic component form a large part of statistical work. Examples include the investigation of the small-sample properties of statistical procedures for which asymptotic theory provides only large sample results and the investigation of complex systems such as those arising in weather modeling. The principles of experimental design described

in the earlier sections of this entry apply to the design of such stochastic simulations.

Anthony C. Atkinson

See also Biochemistry, Contemporary; Biostatistics; Chemical Engineering; Hypothesis Testing; Medicine, 20th Century; Statistical Inference, Bayesian; Statistical Inference, Frequentist; Statistics; Thought Experiments, Scientific and Philosophical

Further Readings

- Atkinson, A. C., Donev, A. N., & Tobias, R. D. (2007). *Optimum experimental designs, with SAS*. Oxford University Press.
- Box, G. E. P., & Draper, N. R. (1969). *Evolutionary operation: A statistical method for process improvement*. Wiley.
- Box, G. E. P., & Draper, N. R. (2007). *Response surfaces, mixtures, and ridge analysis* (2nd ed.). Wiley.
- Cox, D. R., & Reid, N. R. (2000). *The theory of the design of experiments*. Chapman and Hall/CRC Press.
- Dean, A. M., & Voss, D. (2003). *Design and analysis of experiments*. Springer-Verlag.
- Logothetis, N., & Wynn, H. P. (1989). *Quality through design*. Clarendon Press.
- Rosenberger, W. F., & Lachin, J. L. (2016). *Randomization in clinical trials: Theory and practice*, (2nd ed.). Wiley.
- Santner, T. J., Williams, B., & Notz, W. (2018). *The design and analysis of computer experiments* (2nd ed.). Springer-Verlag.

EXPLANATION

The nature of explanation is one of the most enduring problems of philosophy. Explanation has generally been associated with causation, in that when we are explaining an event or phenomenon we are identifying its cause. The ancient Greek philosopher Aristotle's theory of causation can be regarded as an early theory of explanation. At the beginning of the 20th century, the view that was dominantly held was that science cannot provide explanations of natural phenomena.

As the 20th century progressed, however, advances in the philosophy of science meant that a more rigorous independent analysis of the

concept of explanation became necessary. Explanation began to be seen as a primary goal of science. For the first time, new scientific theories about the existence of unobservable entities and processes—for example, atoms, genes, and fields—began to emerge. Traditional empiricists, however, had to resist treating these theories as *true*, because according to their view unobservables were merely posits, and claims about them could not be evaluated as true or false.

In order to resolve this dilemma it became crucial to find a way of characterizing these theories' obvious value while retaining the fundamental principles of rationality deemed central to science. There needed to be a distinction between *truth* and *explanation*.

Truth and Explanation

Theories of explanation can be either realist or epistemic (anti-realist). A *realist* understanding of explanation puts forth the idea that entities or processes posited by an explanation truly exist—and that the explanation is a literal description of an external reality. Realist interpretations hold that *truth* and the explanatory power of a theory involve matters of language which correspond to an actual external reality. This theory of explanation gives us *real* insight into the causal structure of the world.

An *epistemic* interpretation of explanation, in contrast, posits that unobservable entities or processes do not necessarily exist in a literal sense but can provide a useful means of organizing human experience. What we discover during a scientific experiment gives us the means to construct an empirical model—as opposed to furnishing a literal description of reality.

Someone who denies that scientific theories are explanatory in a realist sense does not necessarily deny that they are explanatory in an epistemic sense. And the same is true the other way around, namely, that someone who asserts that scientific theories are explanatory in an epistemic sense may (or may not) claim that they are explanatory in a realist sense. One way of looking at the distinction between truth and explanation is to say that theories that refer to unobservable entities may *explain* the phenomena, but they are not true in a *literal* sense. Put another way, we can say that these theories are true, but they do not in fact *explain* the phenomena.

In these two situations, we can see that the terms *truth* and *explanation* are being used differently in the two statements. In the former, the term *explanation* is being used epistemically and *truth* realistically; in the latter, *explanation* is used realistically and *truth* epistemically. Early 20th-century philosopher and scientist Pierre Duhem (1861–1916) aligned himself with the second interpretation, believing that a physical theory is not an explanation but merely a system of mathematical propositions deduced from a small number of principles representing a set of experimental laws. For Duhem, the basis of explanation involved stripping the reality of the appearance to get to the underlying reality. For him, explanation was the task of metaphysics, not science, since science cannot comprehend reality but exists simply to give order to appearance.

The later rise of analytic philosophy, in particular that of logical positivism, rendered Duhem's classic metaphysical stance unpopular. Logical positivism sought to legitimize philosophical discourse on a similar basis to that employed by the best claims of empirical science. An anti-realist movement—famed for seeking to purge scientific thought of any talk about nature's unobservable aspects, including mechanism, causality, and principles—logical positivism posed science as explanation.

Carl Hempel's Theory of Explanation

Carl Hempel (1905–1997) was a logical positivist, to whom many attribute the first robust theory of scientific explanation for contemporary science. Hempel was concerned solely with logical form and argument—without reference to any actual physical connections between phenomena. *Laws* play a key role in Hempel's concepts of explanation. Hempel's theory used the assumption that laws are *regularities* that meet conditions that are acceptable to empiricists.

In his view, to explain the phenomena in the world of experience was to answer the question “Why?” rather than “What?” This is one of the main objectives of rational inquiry—to go beyond mere descriptions of subject matter by providing an explanation of the phenomenon itself.

For Hempel, answering the question “Why?” did not (as it did for Duhem) involve appealing to a reality beyond experience. His “Why?” expressed

a desire to make predictions about future experiences, and for him the value of a scientific theory rested on its ability to do this.

Hempel proposed two models of explanation: The *deductive-nomological* (DN) model and the *inductive-statistical* (IS) model. Both have the same structure, with premises containing statements of two types: initial conditions C and lawlike generalizations L. The only difference between the two laws is that the laws in a DN explanation are universal generalizations, whereas the laws in IS explanations are statistical generalizations.

An example of a DN explanation containing one initial condition and one lawlike explanation follows:

C: The baby's cells have three copies of chromosome 21.

L: All babies with cells that feature three copies of chromosome 21 will suffer from Down syndrome.

E: The baby has Down syndrome.

An example of an IS explanation:

C: The man's brain was deprived of oxygen for 5 minutes continuously.

L: It is almost always the case that people whose brains are deprived of oxygen for 5 minutes will suffer brain damage.

E: The man has brain damage.

Hempel preferred DN explanations to IS explanations, because the deductive relationship between the premises and conclusion maximized the predictive value of the explanation. For him, IS arguments were explanatory only insofar as they approximated DN explanations by giving them a high degree of probability of explaining an event.

Hempel saw explanation as a concept that should be understood in terms of logic. A good DN explanation relies on there being true premises. But in order to qualify as a DN explanation, an argument needs simply to exhibit the DN structure. This illustrated Hempel's logical positivist position, which required that all epistemically significant scientific concepts should be analyzed in scientific terms.

When it comes to IS (based on statistics) explanations, however, their inductive character can always be undermined by the addition of new information. For example, in the earlier IS explanation, the case of

the probability of a man suffering brain damage may be reduced by the fact that he was submerged in a cold body of water. A proposed IS explanation, then, even if the premises are true, would fail to predict the fact and would have no explanatory significance.

Criticisms of Hempel's Theory

Hempel's lack of enthusiasm for statistical explanation was criticized by many modern scientists who saw its use for explanation as indispensable. This was because of the counterintuitive implication that for inherently low probabilities of events, no explanations are possible.

For example, it may not hold statistically that smoking three packs of cigarettes a day for 30 years will lead to a person's contracting lung cancer. If we follow Hempel's theory, it follows that a statistical law about smoking cannot be used to explain the occurrence of lung cancer, a position which is at odds with what we know about smoking and lung cancer.

Hempel might counter this by stating that when our theories do not allow us to predict a phenomenon with a high degree of accuracy, it is due to our incomplete knowledge of the conditions. However, this still puts us in the uncomfortable position of having to base a theory of explanation on the dubious metaphysical position that all events must have determinate causes.

Hempel's view was also criticized because DN arguments that have true premises may be unexplanatory. According to Wesley Salmon, the relevance of this problem is as follows:

Ci: Adam takes birth control pills.

Cii: Adam is a man.

L: Men who take birth control pills do not become pregnant.

E: Adam does not become pregnant.

While this qualifies as explanatory according to Hempel's theory, the premises are obviously explanatorily irrelevant to the conclusion.

Sylvain Bromberger (1924–2018) took Hempel to task on the question of asymmetry: On the one hand, Hempel's model can explain the period of a pendulum in terms of the length of the pendulum together with the law of simple periodic motion; simultaneously we can explain the length of the

pendulum in terms of its period in accord with the same law of motion. Intuitively, we may feel that the first is the proper explanation, while the second is not. We can make the same point in the following way:

C: The barometer is falling rapidly

L: Whenever the barometer falls rapidly, a storm ensues

E: A storm is imminent

Despite the fact that the falling barometer indicates the storm's approach, it is counterintuitive to state that the barometer *explains* the approach of a storm. It is instead the approaching storm that *explains* the falling barometer.

These problems of relevance and symmetry reveal how difficult it is to present a theory of explanation that has no relation to causality. Hempel cannot have causal relations as an option, since causation is at the top of the list for the anti-realist's view of metaphysically suspect concepts. Involving causality would also undermine the view that explanation can be understood in terms of an epistemic, rather than a metaphysical, relationship. Hempel responded by stating that whether explanation has a practical value is not something that can be determined via philosophical analysis. This reply, although adequate, did not satisfy his critics.

Lawlike and Accidental Generalizations

Hempel's model of explanation requires there to be at least one lawlike generalization. This also presents a problem for the DN model. Hempel distinguished between lawlike generalizations and accidental generalizations, the latter being those that might be true but not as a result of a law of nature. For example, when saying "All my shirts need ironing," we are stating an accidental fact rather than an event that is subsumed under the laws of nature.

This makes it doubtful whether Hempel's epistemic view of explanation can be considered consistent, since there is no epistemically sound way of distinguishing between lawlike generalizations and accidental generalizations. No finite number of observations can back up the claim that a natural regularity is due to a natural necessity. In the

absence of such a criterion, Hempel's model does not appear to fit with an epistemic view of explanation, and that explanation can be understood in purely logical terms.

Contemporary Explanation Theory

With the decline of logical positivism and the increasing success of modern theoretical science, arguments about whether unobservable entities existed became redundant. Explanation as attributable to a cause came back into fashion.

Wesley Salmon's (1925–2001) influential realist causal account of explanation takes the view that real processes and entities are conceptually necessary in order to understand why explanations work. He rejected all epistemic theories of explanation as incapable of being able to bring about adequate scientific understanding.

For Salmon, scientific understanding is not just concerned with being able to hold justified beliefs about the future. He argued that even someone with a near-perfect knowledge of all natural laws and universal conditions and in possession of perfect predictive knowledge would potentially still lack scientific understanding. Having predictive knowledge does not entail having awareness of concepts such as causal relevance that are essential to allow insights into true causal processes and pseudo-processes.

Salmon's causal realism rejects Hume's influential concept of causation which imagines a series of linked chains of events. Instead he sees causation as a continuous series of causal processes and interactions or set of instructions for producing an explanation of a particular event or phenomenon. By starting with a list of statistically relevant factors, then analyzing these factors we end up with a causal model of statistical relationships, which can be empirically tested in order to select those which best support the evidence. This is a *reductionist theory* which rests on finding the causal mechanisms of a phenomenon. Salmon's theory does not lend itself easily to explanations of mental events however which do not appear to be easily reducible to a set of causal relationships.

Like Hempel, Salmon's theory was an ideal or higher form of explanation theory, and these kinds of theories of explanation began to be criticized for failing to resonate with scientists or ordinary

people in their everyday lives. Another approach to explanation involved the relationship, beliefs, and intentions used by people on a day-to-day basis.

Explanation as a Form of Communication

As well as these theoretical and essentially scientific approaches to explanation theory, certain philosophers have grounded their understanding of explanation in the way in which people perform the act of explanation. This so-called *ordinary language philosophy* focuses on the linguistic and communicative aspects of explanation. This involves answering questions with the aim of furthering the understanding between individuals.

Peter Achinstein developed an *illocutionary* theory of explanation, which defined the act of explanation as the attempt by one individual to produce understanding in another by answering certain kind of questions in certain ways. He rejected Salmon's narrow interpretation of explanation and its association with causation. According to Achinstein, we attempt to gain understanding by asking different questions such as "Who?" "What?" "When?" and "Where?" in order to find explanations.

Whereas traditional theories of explanation attempted to understand the *logic* of explanation, Achinstein tried to describe the *process of explanations*. For example, when a person asks for an explanation of "why the lights in the house have gone out," they implicitly formulize their question in such a way so that they will receive a reply that will enable them to turn them back on again. The explanation they require must practically help them put the lights back on. Receiving an answer in terms of, say, mathematical equations would be of no help to them—but must be provided in a practical understandable manner.

Achinstein ruled out tautological answers as explanations too, for example, "Mr. Smith died because he ceased to live" as being of no explanatory use. He also counted as nonexplanatory, non-relevant, and noncontextual answers to scientific questions. For example, if we ask, "What is the speed of light in a vacuum?" and receive the answer "186,000 miles per second" we have not been given a satisfactory explanatory answer but simply a large number with no comparative reference to velocities that mean anything to us.

Naturally this does not mean that the speed of light cannot be explained—it's just that the answer needs to be framed in terms that we can relate to, for example, "light can travel around the earth eight times a second."

Achinstein has been criticized for being somewhat vague when it comes to exactly what are *content-giving* propositions. He failed to narrow down a list of questions which would qualify as requests for explanations, making it difficult to identify what it is specifically about any act of explanation that enables it to produce understanding. It's also difficult to infer that simply asking for a certain kind of knowledge will actually lead to achieving the desired knowledge.

Cognitive Science and Explanation

Another departure from standard relational analysis of explanation can be found in cognitive science. A cognitive approach to explanation maintains that an explanation is a kind of mental representation that results from or aids this understanding. There are divisions in the field of cognitive science as to whether explanation is conceived as the process and result of belief revision—or whether it is the activation of patterns in a neural network that provides the basis of explanation.

The latter perceives the act of explanation as a purely cognitive activity and explanation as a mental representation resulting from or aiding this activity. Explanation is thereby considered a universal phenomenon which can be conscious and deliberate or unconscious and instinctive, requiring no explicit propositional knowledge.

For example, a mother hearing a high-pitched cry coming from her child's room will immediately rush to that room to come to their aid. This act could be described as instinctive or based on inference, but whatever the reason we can say it was as a result of the mother having *explained* to herself the crying noise as that of the crying of their child. This gives explanation a theoretical rather than a logical or metaphysical status.

Mental Models

One approach used by cognitive scientists incorporates both traditional belief-desire aspects and neuroscience resulting in the idea of a mental

model. The term was first put forward in 1943 by Kenneth Craik in his book *The Nature of Explanation*. In his book, Craik suggested that the mind constructs *small-scale* models of reality that it uses to anticipate events. Mental models arise from mental representations that result from the activation of an aspect of a network of condition-action (if-then)-type rules.

When clustered together, these conditions (rules) apply to get action results. They rest on the assumption that reasoning depends not on logical form as used in formal rule theories but on models constructed from perception and imagination, somewhat akin to a physicist's diagram in that the structure is analogous to the structure of the situation that it represents.

Based on a small set of fundamental axioms or assumptions, each mental model represents a possibility, capturing what is common to all the various ways in which the possibility may occur. Mental models are based on a principle of truth, since all the possibilities represent what is true according to a proposition. Individuals infer that a conclusion is valid if it holds in all the possibilities, and counterexamples act to refute valid inferences. Critics of mental model theory include those who cite deduction-based theories based on inference rules rather than relying on models.

Naturalism and Explanation

Naturalism is associated with the rejection of any kind of explanation of natural phenomena that refers to unnatural phenomena. Many naturalists embrace the idea of inference as the best form of explanation, as a fundamental principle of scientific reasoning. They hold that if we treat the existence of unobservable entities as a scientific hypothesis, then this can provide an explanation of the success of the theories that employ them. In contrast anti-realism cannot provide such an explanation since they view theories that refer to unobservable entities as literally untrue and the success of scientific theories unknowable. As mentioned at the start, traditional realists, unlike scientific realists, do not think of realism as a scientific hypothesis but as an independent metaphysical thesis.

The State of Explanation Theory Today

Our knowledge of the nature of explanation has improved since its beginnings, with each new model giving us pause for thought. Whether it is possible to construct a single model of explanation to fit all areas is a moot subject, since the types of explanatory goals—the sorts of explanatory information sought and what is thought to be achievable, testable, or discoverable—vary so widely across disciplines. In future, it is predicted that models of explanation will be developed that are sensitive to disciplinary differences.

Ideally, these models should also reveal commonalities across disciplines as well as allow us to see why explanatory practice varies along with its significance. For example, biologists often describe their explanatory goals as the discovery of mechanisms, while physicists talk about explanatory goals in terms of laws. While this may just be a purely terminological difference, it may also be the case that we need to uncover the background to what a mechanism is, as understood by a biologist, and how this can impact on a theory of explanation.

Thora L. Fitzpatrick

See also Deduction; Evidence; Induction; Inference; Mental Models

Further Readings

- Achinstein, P. (1983). *Nature of explanation*. Oxford University Press.
- Bromberger, S. (1966). Why Questions. In R. G. Colodney (Ed.), *Mind and cosmos*. University of Pittsburgh Press.
- Craik, K. J. W. (1943). *The nature of explanation*. Cambridge University Press.
- Duhem, P. (1991). *The aim and structure of physical theory*. Princeton University Press. (Original work published 1954)
- Friedman, M. (1974). Explanation and scientific understanding. *Journal of Philosophy*, 71, 5–19.
- Hempel, C. G. (1962). Deductive-nomological vs. statistical explanation. In H. Feigl & G. Maxwell (Eds.), *Minnesota studies in the philosophy of science* (Vol. III, pp. 98–169). University of Minnesota Press.
- Hempel, C. G., & Oppenheim, P. (1948). Studies in the logic of explanation. *Philosophy of Science*, 15, 135–175. <https://doi.org/10.1086/286983>
- Johnson-Laird, P. N. (1983). *Mental models: Towards a cognitive science of language, inference and consciousness*. Cambridge University Press.
- Salmon, W. C. (1984). *Scientific explanation and the causal structure of the world*. Princeton University Press.
- Van Fraassen, B. (1980). *The scientific image*. Oxford University Press.

F

FACT VERSUS THEORY

The relationship between *fact* and *theory* is a matter of controversy in science and in the philosophy of science. Such controversies stem from the debate between different varieties of realism/scientism and antirealism/antiscientism and their different ontological and epistemological assumptions. A first controversy arises as to whether the use of facts as conditions, criteria, and rules for establishing justified true belief (i.e., a belief's epistemic value) is or is not a distinctive aspect of scientific theories regarding other kinds of theorizing (e.g., common knowledge, history, journalism).

In this vein, antiscientism (e.g., Paul Feyerabend) asserts that appeals to facts—a distinctive feature of scientific theories—are no more than *rhetorical bullying* (i.e., *nominal identity* between scientific theories and other kinds of theorizing). Susan Haack's *critical commonsensism* supports the thesis that science is not radically different from other kinds of inquiry in the use of evidence: Epistemic values, chief among them respect for facts, are also relevant to other forms of theorizing (i.e., *difference of degree* between scientific theories and other forms of theorizing). Finally, scientism (e.g., Alexander Rosenberg) postulates that the very meaning of scientific inquiry is to build and control theories on factual entities inaccessible to common knowledge and other forms of theorizing (i.e., *nominal difference* between scientific theories and other forms of theorizing).

A second controversy on the relationship between fact and theory arises from the very definition of fact—particularly whether a fact is considered equivalent to a *phenomenon*. The different answers to this question can be synthesized in the debate between phenomenalism (as a variety of antirealism) and scientific realism (as a variety of realism). Phenomenists have defined facts as the collection of those entities that are the object of experience, that is, *phenomena*. In a different manner, scientific realists look for facts behind phenomena (or appearances), which they postulate are accessed through reason. For scientific realism, a phenomenon is the class of *observable* facts that is characterized by being an event accessed through the senses, which, together with the class of *unobservable* facts accessed through reason or nonsensory intuition, constitute reality. In short, for scientific realism, *phenomenon* and *fact* are not synonymous; a phenomenon is only a class of fact. For phenomenalism, fact and phenomenon are equivalent; facts consist of phenomena.

A third controversy on the relationship between fact and theory results from the definition of theory, specifically of scientific theories. For scientific realism, given its ontological assumptions, any project of scientific theory that is not trivial must have the objective of discovering what lies beneath what is observable. Observable facts are real but only superficial. Because they are superficial, they must be explained in terms of unobservable facts that are capable of being scrutinized, and they

must be scrutinized via scientific theories. For phenomenalism, given that reality is accessed inevitably through the intersection between subject and object, the world is dependent on the subject; the scope of scientific theories is circumscribed inherently to those facts to which we have access through experience. Thus, scientific theories are conceived as giving descriptions of observable entities in a subject-dependent world.

A fourth controversy emerges on the intertwining of theory and fact in the production of useful scientific results. Two questions stem from this debate. First, to what extent do scientific theories represent and explain facts? Second, what place do facts have in testing scientific theories?

Regarding the first question, scientific realism has affirmed that scientific theories can produce (fallible) objective truths about at least some facts (epistemological realism). Antirealist traditions, with some variations among their theses, deny the possibility of objective truth (epistemological skepticism). Hence, one variety of antirealism, postmodernism (e.g., Jean-François Lyotard), proposes that scientific theories are *metaphors* rather than representations of facts. Another variety of antirealism, constructivist relativism (e.g., Bruno Latour and Stephen Woolgar), posits that science is a form of fiction or discourse equivalent to any other, such as a literary text. Other alternative to scientific realism is constructive empiricism (e.g., Bas van Fraassen). Constructive empiricism is a contemporary—and influential—variety of phenomenalism, which postulates that science aims to produce theories that are *empirically adequate* rather than true, where a theory is empirically adequate if what this theory affirms regarding observable phenomena (i.e., those entities that can be directly observed by the human eye) is true.

Regarding the second question, a variety of realism, naive realism, postulates that the truth values of scientific theories derive from their testing with facts. By contrast, other philosophers—such as Norwood Russell Hanson, Thomas Kuhn, and Stephen Toulmin—have argued that because theory is a tribunal of what is considered fact, facts cannot be used to test a theory without committing to that very theory (the *problem of circularity*).

Ontological Status of Fact and Epistemological Status of Theory

The epistemological status that we assign to scientific theories depends on how we conceive of facts' ontological status. As Mario Bunge maintained, epistemology is the result of ontology: depending on what is believed to be in the world, what is believed about what it is possible to know and how to know it are derived.

Regarding the notion of fact, a fundamental difference emerges between realism and antirealism on how *properties* of facts are conceived ontologically. Following the Kantian distinction between *noumena* (or things-in-themselves) and *phenomena* (or things-for-us or things as perceived), scientific realism holds that facts have two types of properties: primary and secondary. Primary properties are objective or subject-independent, and secondary properties are subjective or subject-dependent. For example, the wave frequency of sound is a primary property because it is found in the world (i.e., it is subject-independent), whereas the tone is a secondary property because it emerges in the ear–brain system (i.e., it is subject-dependent).

By contrast, according to varieties of antirealism, primary or objective properties would be derived, imaginary, or (in the best of cases) inaccessible to knowledge. Thus, classical varieties of antirealism either deny the independent existence of the physical world and affirm that only the mind and its ideas exist (George Berkeley's subjective idealism) or do not admit more reality than that of experience, understanding by such the set of subjective perceptions (Edmund Husserl's ontological phenomenalism).

In particular, phenomenalism casts doubt on whether we can access something unobservable. Phenomenalists reduce the world to what Ernst Mach called *a complex of sensations* and claim scientific theories describe only observable facts. To a large extent, Rudolf Carnap's unsuccessful attempt to eliminate theoretical terms and leave only observational terms derived from the phenomenalist ontology—and from the antimetaphysical program—that characterized logical empiricism (or logical positivism or neopositivism: these labels are used interchangeably here). In short, and according to phenomenalists, the world

would be reduced to secondary properties (ontological phenomenalism), and sensible human beings would get to know the world through common experience (epistemological phenomenalism). Denying the existence or accessibility of the primary or objective properties, scientific theories must be directed to the study of the secondary properties or phenomena.

In other words, scientific realists recognize the ontological and epistemological statuses of primary properties and secondary properties, whereas phenomenalists recognize only the ontological and epistemological statuses of secondary properties (which the latter call *experience* or *appearances*). In this way, phenomenalism maintains the following: (a) The properties of facts are circumscribed to those of sensible experience (i.e., secondary properties for scientific realism) and (b) if there were primary properties, then we could not access them because inherently we know the world through the subject–object interface. By contrast, scientific realism states the following: (a) Facts have both primary and secondary properties (the former are found in the world, and the latter are found in the brain), and (b) we can access the primary properties through reason, nonsensory intuition, and especially scientific theories.

Nevertheless, for scientific realists, the ontological and epistemological statuses of facts' primary and secondary properties are more complex. In ontological terms, scientific realism claims there are no secondary properties in the physical world. There are no colors, smells, or tastes in the external world, because they are located in the subject–object interface. Thus, all secondary properties require two sources: the brain and something else in the world. Hence, although some primary properties (e.g., chemical composition) are intrinsic, all secondary properties (e.g., taste or smell) are relational; they are all only for us, and none exists in itself. Consequently, the ontology of most secondary properties of facts requires the inclusion of the brain. Without brains, most secondary properties would have no existence as such.

In epistemological terms, scientific realism asserts that scientific theories do not contain concepts denoting secondary properties—unless secondary properties are their object of study, as is the case with psychology. Scientific theories (e.g., of physics, chemistry, biology) describe noumena

(i.e., physical things), not phenomena (i.e., mental processes)—unless mental processes are their object of study, as is the case with neuroscience. There is no trace of phenomena (appearances to a sensitive being) or secondary properties in the formulas of physical theory. No scientific theory contains variables denoting secondary properties (colors, for instance).

In addition, scientific realists have proposed that scientific theories cannot explain secondary properties (e.g., the taste of food) without primary properties (e.g., the chemical composition of food). For example, it is impossible to explain a secondary property of lemons, such as their acid taste—which our taste buds recognize—without taking into account a primary property of lemons—that is, that they contain hydrogen ions.

Fact, Theory, and Epistemic Justification

A fundamental nexus between fact and theory is related to the problem of the epistemic value of theory. The epistemic value of a theory involves the central problem of when and how we are justified in holding that theory to be true. Regarding fact, a question that emerges about the epistemic value of a theory is whether, how, and to what extent facts justify its truth. This question is considered particularly meaningful in assessing the epistemic value of scientific theories.

Although the postulated epistemic relationship between scientific theory and fact varies according to different philosophies of science, a long tradition has maintained that a theory is considered scientific when it proposes a certain set of refutable hypotheses—that is, when a theory seeks to confirm or disconfirm its adequacy to the facts. The correspondence between fact and theory has been considered by both verificationism (i.e., a given theory is scientific if it entails verifiable observable consequences) and falsificationism (i.e., a given theory is scientific if it has falsifiable observable consequences) as a meaningful criterion for the epistemic justification of scientific theories and as a demarcation criterion between the scientific and the unscientific. Science systematically seeks to determine whether facts confirm or falsify theories. Thus, the use of empirical evidence (i.e., facts) distinguishes science from other, nonscientific, cognitive endeavors.

However, more contemporary philosophers of science have emphasized that the links between fact or the factual state of affairs (i.e., nonconceptual entities) and theory (i.e., a conceptual entity) are more complex. First, sentences, not facts, confirm or disconfirm theories. One important consequence of the *linguistic turn* in philosophy—frequently associated with analytic philosophy—has been to concentrate a theory's epistemic value on the logic of observation rather than on objects or facts. Hence, a scientific theory is conceived as a system of sentences or sentence-like structures (e.g., propositions, statements, claims) to be tested by comparison with observational evidence. Consequently, it follows that theories are tested not against facts but against sentences or propositions on observed things.

Second, the problem of the relationship between fact and theory makes sense in the factual sciences, not in the formal sciences. In this vein, some approaches fail to distinguish between two different sorts of truth, that is, formal truth (proper to the formal sciences) and factual truth (typical of the factual sciences). On the one hand, a formal science is by definition a discipline that contains only formal propositions (e.g., mathematics, logic). On the other hand, a factual science is by definition a discipline that contains at least some factual propositions; they are the sciences that describe, explain, or predict things or events that belong to the real world, natural or social (e.g., natural sciences, social sciences).

Therefore, in the formal sciences, (a) objects (e.g., mathematical and logical entities) are defined in a purely conceptual way without resorting to any factual or empirical means, and (b) formal proofs and refutations (e.g., logical and mathematical theorems) are strictly conceptual processes that do not make any reference to empirical entities or facts. Consequently, in the formal sciences, theory is epistemically independent of fact because formal truth depends only on logical rules of inference. Thus, an epistemic relationship between fact and theory makes sense only in the factual sciences.

Third, confirmation theory—simply defined as the study of the logic by which scientific theories may be confirmed or disconfirmed by evidence or facts—has proven a rather difficult endeavor. The ideas of many philosophers of science, such as

Pierre Duhem and Willard Van Orman Quine (and, more recently, for the Bayesian confirmation theories), have challenged the venerated confirmation theory of Carl Gustav Hempel. In particular, the governing idea in traditional confirmation theory, according to which each statement taken in isolation can admit confirmation or disconfirmation, is questioned as a form of reductionism. This idea has been criticized on two main grounds. First, because empirical statements are interconnected, they cannot be disconfirmed singly (the Duhem–Quine nonseparability thesis). Second, and as a corollary, empirical evidence tests a theory as a *corporate body*. As a consequence, if a theory is falsified, then it is impossible—at least in some sciences—to investigate where the error lies (the Duhem–Quine nonfalsifiability thesis).

Fourth, an additional difficulty for the epistemic value of fact versus theory emerges from the idea that observation is permeated inherently with theory, which thus has weakened the fundamental distinction between theoretical terms and observational terms. Observational terms garnered special epistemic status because they seemed more direct and therefore less distorted than theoretical terms. In physics, the conception according to which experiments show observation to be completely permeated by theoretical terms has gained acceptance: Observation is theory-laden to the point where, as Pierre Duhem observed, it becomes impossible to express fact in isolation from theory. In neuroscience, it has been increasingly difficult (if not impossible) to reconcile the idea of highly processed data such as functional magnetic resonance imaging pictures with the search for a pure observational language. The consequence of those developments in the factual sciences is an increased blurring of the distinction between theoretical terms and observational terms, which makes it difficult to accept both that a theory can be tested with observational data and that it can be asserted that Theory A accounts for the facts better than does Theory B.

Fifth, other philosophers of science—such as Thomas Kuhn—have argued that the meanings of terms are determined by a certain theory or a *paradigm* to the point where even the same terms cannot be compared in various theoretical contexts because their meanings depend intimately on, and vary with, each theoretical context. For

example, because Newton's theory is based on a paradigm different from that of Einstein's theory (according to Kuhn), terms such as *mass*, *space*, and *time* have different meanings per each theory. An epistemic consequence of the Kuhnian approach is that it calls into question the idea that it is possible to judge each theory against the other in light of observed facts.

Finally, some philosophers, such as Stephen Toulmin, have raised the problem that facts (i.e., their observation) are not accessed directly but through *ideals* or *ideals of natural order*, that is, through theories that establish what are considered *facts*. Consequently, the theories themselves establish the facts against which the theories are subsequently tested. In the factual sciences, the theory is not related to the facts merely for their confirmation or disconfirmation; rather, the theory provides the framework for what are regarded as facts. In this sense, the claim is that none of the factual sciences apprehends its object as it is but as its own analytical operations modify that object. For this reason, the factual sciences are said to transcend facts—they discard some and produce new ones. Furthermore, in the factual sciences, theory is a tribunal of what is regarded as fact.

Accordingly, the relationship between theory and fact is twofold. On the one hand, the facts are used to confirm or disconfirm a theory; on the other hand, the theory establishes those parts of reality that are considered facts. Last, because of the paradoxical relationship between fact and theory, some philosophers of science—such as Karl R. Popper, Norwood Russell Hanson, and, more recently, Allan Franklin—have worried that drawing upon theoretical assumptions from the very theory to be appraised yields vicious circularity.

Rodolfo Sarsfield

See also Epistemology; Justification; Philosophy of Science; Warrant

Further Readings

- Ayer, A. J. (1953). *Language, truth, and logic*. Gollancz.
 Bunge, M. (2007). *Chasing reality: Strife over realism*. University of Toronto Press.
 Carnap, R. (1932/1934). *The unity of science*. Kegan Paul. (Original work published 1931)

- Duhem, P. (1991). *The aim and structure of physical theory* (Philip P. Wiener, English Trans.). Princeton University Press. (Original work published 1954)
 Feyerabend, P. (1975). *Against method: Outline of an anarchistic theory of knowledge*. Humanities Press.
 Haack, S. (2003). *Defending science within reason: Between scientism and cynicism*. Prometheus Books.
 Hanson, N. R. (1958). *Patterns of discovery: An inquiry into the conceptual foundations of science*. Cambridge University Press.
 Hempel, C. G. (1945). Studies in the logic of confirmation. *Mind*, 54(214), 97–121. <https://doi.org/10.1093/mind/LIV.214.97>
 Kuhn, T. S. (1970). *The structure of scientific revolutions*. University of Chicago Press.
 Laudan, L. (1983). The demise of the demarcation problem. In R. S. Cohen & L. Laudan (Eds.), *Physics, philosophy and psychoanalysis* (Boston Studies in the Philosophy of Science, Vol. 76, pp. 111–127). Springer. https://doi.org/10.1007/978-94-009-7055-7_6
 Nagel, E., Suppes, P., & Tarki, A. (Eds.). (1962). *Logic, methodology and philosophy of science: Proceedings of the 1960 international congress*. Stanford University Press.
 Neurath, O. (1983). *Philosophical papers 1913–1946*. Reidel.
 Popper, K. R. (1959). *The logic of scientific discovery*. Hutchinson & Co. (Original work published in German by Springer, 1935)
 Quine, W. V. (1951). Main trends in recent philosophy: Two dogmas of empiricism. *The Philosophical Review*, 60(1), 20–43. <https://doi.org/10.2307/2181906>
 Rescher N. (1997). *Objectivity*. University of Notre Dame Press.
 Rorty, R. (1991). *Objectivity, relativism, and truth*. Cambridge University Press.
 Toulmin, S. E. (1963). *Foresight and understanding*. Harper Torchbooks.
 van Fraassen, B. C. (2002). *The empirical stance*. Yale University Press.

FALSIFIABILITY

Falsifiability focuses on theory disconfirmation using deduction, and application of this epistemology and method signifies a science, according to the tenets of critical rationalism, formulated by Karl Popper (1902–1994), a philosopher of science. Popper received much

acclaim during his lifetime, although many recently published philosophers have distanced themselves from his ideas. In contrast, most contemporary scientists and social scientists with a research focus are strongly influenced by his thoughts on scientific methods, including falsifiability. Before presenting the concept, some information about the philosopher is provided, since biographic factors played a role in Popper's theoretical formulations.

Karl Popper

The philosopher was born and educated in Vienna, Austria. In later recollections of these early years, Popper recalled Albert Einstein's (1879–1955) Vienna lecture of 1921, which he attended at the age of 19. Einstein's discussion of the theory of relativity resonated with Popper, especially since the British astronomer and physicist Arthur Eddington (1882–1944) had led an expedition to Principe, a West African island, to observe the 1919 solar eclipse, vindicating the theory, although the expedition also had potential to negate the theory. This final point was one of many stimulants suggesting to Popper that falsifiability rather than testability was the correct way to move scientific endeavors forward.

Other important links in Popper's philosophical path toward his later views of science were his early interactions with admirers of Alfred Adler (1870–1937) and Sigmund Freud (1856–1939), both medical doctors, active in Vienna, and founders of psychoanalysis. Popper had volunteered with disadvantaged children, and in the process, took a patient study to Adler, who managed to slot the case into his theory in a process that Popper later considered to be intellectually dishonest.

In contrast to his later reevaluation of the psychoanalysts, Popper maintained greater respect for a group of psychologists involved in the Würzburg school, based in Würzburg, Germany, specifically Oswald Külpe (1862–1915) and Otto Selz (1881–1943). Selz instructed Popper, who later incorporated his teacher's thesis that problem-solving is essential to processing ideas into his definition of science.

A fourth key event during Popper's youthful years was his involvement with Marxism, before Austrian government changes engendered his

realization of how political extremism curtailed freedom and destroyed lives. Popper's parents had converted to Protestantism from Judaism, and in the anti-Semitic climate of Austria, leading up to the Nazi takeover of Europe, and World War II, his family's Jewish background became significant. Indeed, the Nazis killed many of his relatives, although Popper managed to leave Austria and was spared.

Science

Popper's 1934 book *Logik der Forschung* remained inaccessible to English speakers until 1959, when *The Logic of Scientific Discovery* was released. The English version was an unamended translation of the earlier publication, with appendices and footnotes augmenting the original book. Popper favored a clear, succinct manner of writing, and for the most part, his work is accessible to scientists and social scientists, not just philosophers.

The Logic of Scientific Discovery focused on how science should move forward, both epistemologically and methodologically. According to this perspective, science is not static but constantly progressing as theories are falsified and alternatives proposed in their place. Popper was not presenting the history of scientific advances as a descriptive exercise but rather prescribing how science should proceed, with a strong empirical basis. He presented a problem-oriented science based on deductive reasoning.

In this book, Popper reviewed his earlier fascination with Marxism and psychoanalysis, critiquing their methods as unscientific. In contrast, he pointed to Albert Einstein's work on gravity, with its risks of falsification by the 1919 solar eclipse data.

Popper was interested in the context of justification, specifically how theories are explained and then examined through comparisons with the real world. He did not deal with the context of discovery, which covers how theories are created, and what inspiration or thinking process underlies their formulation.

He demarcated a clear boundary between scientific and nonscientific theories and claimed that the limits of science were determined by the use of falsification. This means that scientists must

consider how to disprove their theories and attempt to the best of their abilities to apply empirical evidence to this end. Popper used the analogy of a net in his discussion of theory, with the throwing of the theoretical net into the world signifying the attempt to understand reality. As science advances, the net mesh size is gradually narrowed, meaning that the world is comprehended in greater detail, with continual skepticism as the procedure continues. However, there is never certainty to this process, since a theory is only accepted so long as it has not been disconfirmed, and there is always the potential of future disconfirmation.

In contrast, positivists, such as the members of the Vienna Circle encountered in Popper's youth, believed that the weight of accumulated positive tests of a theory provided greater justification for the theory being correct. Maintaining the image of the net, such an approach would thicken the net's mesh, rendering it ever more improbable that disconfirmation could ever occur, and supporting a more static scientific universe.

Popper's ideas about demarcating the boundary between science and nonscience were challenged by Imre Lakatos (1922–1974), a colleague at the London School of Economics. Their initially friendly relationship was marred when Lakatos published increasingly radical amendments of Popper's ideas. Lakatos felt that Popper's criterion of falsifiability as key for a science was too inflexible, and he suggested that various degree of falsification might occur in science, as practiced by actual scientists. He also suggested that there might be higher degrees of falsification during some time periods, with lesser gradations at others.

Using Popper's metaphor of the net of science, including the key element of falsification, it is apparent that according to his reasoning, psychoanalysis is not scientific because there is no evidence that can be presented to prove it incorrect. However, the discipline may be considered a pseudoscience, since with some modifications, such as the addition of falsification to examine it, there is potential for it to become scientific. In contrast, the cases of astrology, phrenology, and other forms of fortune telling are simply unscientific, since there is no way to create tests to disconfirm them. They depend on belief. In order to be a science, a field of

study must use a scientific method, incorporating falsification and deduction.

Popper's use of falsification as key to defining a science has received some pushback from academics and support by the courts. In 1981, Judge William Overton ruled in an Arkansas case that falsifiability, as understood by Popper, was one of the four key pillars to determining a science. This meant that creationism, the belief that the universe and life originated from supernatural acts of divine creation, as in the biblical account, did not have to be taught in schools' biology classes. Within 6 years, a Louisiana case was ruled similarly in the Supreme Court. Falsification also appeared in those court documents. Michael Ruse, a philosopher, wrote about both cases.

In an interesting twist outside the court, recent examinations of astrology by scientists have used falsification as a way to disprove its scientific validity. However, semantically, for Popper's definition of a pseudoscience, this very application of deduction moves astrology into the category of a science, as clearly presented by Hansson. This suggests that some reconsideration of Popper's definitions may be overdue, which moves the discussion into the next topic.

Scientific Method

As mentioned previously, Karl Popper presented his prescriptive scientific method in *The Logic of Scientific Discovery*. He focused on the investigative methods of research, rather than the theory formation portion of the exercise. While positivists encouraged scientists to confirm hypotheses and theories, Popper moved in the opposite direction, upholding the necessity of falsification for a true science and specifying the method to be used.

In his treatise on method, Popper cited the work of the astronomers Johannes Kepler (1571–1630) and Galileo Galilei (1564–1642). This was not an attempt to recount the history of science; instead, he was promoting the idea that his ideas about science, and its dependence on falsifiability, are demonstrated in the reasoning of Kepler and Galileo (using his first name by convention) when they disconfirmed the suggestion that the sun revolved around the earth, thus supporting a heliocentric theory of the

universe. Popper also revisited the writings of Xenophanes (ca. 540–478 B.C.E.), a Greek philosopher, and claimed some support from this work for his methods.

Only a few years later, in 1962, Thomas Kuhn (1922–1966) presented *The Structure of Scientific Revolutions*, which also covered the work of Kepler and Galileo, although with a different focus. Kuhn argued that scientific revolutions occur with paradigm changes involving the sociology of the field, rather than according to the scientific merits of theories.

Popper, however, was completely uninterested in sociological interpretations of theory development and focused instead on the type of language applicable to the presentation of hypotheses, terming these *universal statements*. In some of his examples, he also allowed the substitution of mathematical symbols for words, based on a similar formatting style. Universal statements were examined by deductive reasoning to create means for disconfirming hypotheses, rendering the method scientific.

Deduction is working from universal statements to suggest a prediction, which can be experimentally examined with real-world data, and will serve to disconfirm the hypothesis. In contrast, *induction* is working from observations to a logical statement of the observations, often considered to be the truth of an event when presented by an observer. For example, if someone watches waves breaking on the shore, then it would be possible to formulate a theory about the size of the waves and their speed to calculate the amount of shoreline deposition or erosion. A theory based on these data, however, is difficult to disconfirm using the same information.

Despite Popper's focus on deduction, it is clear that many scientists and social scientists use a combination of induction and deduction as they apply the scientific method to real-world problems. For example, an archaeologist might use ideas about how human settlements function, based on debris patterns in one part of the world to suggest testable hypotheses about ceramic deposition at a site in another region. These hypotheses would be examined with archaeological excavation and analysis, in a process using induction, deduction, and testable hypotheses, along with falsification attempts.

To encourage well-founded attempts at falsification, Popper encouraged the incorporation of

high empirical content in theories. Enhanced falsifiability options occurred with theories pertinent to more real-world scenarios, since this would provide more opportunities to examine them with data. In addition, greater exactitude in predictions also improved the falsification exercise.

Popper also wrote about the importance of intersubjective testability, namely that all observers of the phenomena should be able to record the same results. In order for a theory to be disconfirmed it needs to be replicably falsified. In theory, falsification can occur with one negative instance, but Popper suggested that falsification should be repeatable under experimental conditions. This allows for falsification errors based on equipment failures or experimenter inexperience.

In conclusion, given the specificity of Popper's ideas and his somewhat rigid methodological stance, it is perhaps not surprising that philosophers have challenged his theory, while scientific experimenters outside philosophy have been more accepting of the broad sweep of his ideas. More recently, Popper's demarcation between science and pseudoscience has played an influential role in the legal and academic battles to remove creationism from public school curricula and has moved his terminology into a new arena. Falsifiability is likely to remain an important concept in the coming years.

Susan J. Wurtzburg

See also Deduction; Epistemology; Hypothesis Testing; Induction; Philosophy of Science; Pseudoscience; Theory Change

Further Readings

- Bird, A. (2018). Thomas Kuhn. In E. Zalta (Ed.), *The Stanford encyclopedia of philosophy* (Winter ed.). Retrieved November 2020, from <http://plato.stanford.edu/archives/win2018/entries/thomas-kuhn/>
- Hansson, S. O. (2017). Science and pseudo-science. In E. Zalta (Ed.), *The Stanford encyclopedia of philosophy* (Summer ed.). Retrieved November 2020, from <http://plato.stanford.edu/archives/sum2017/entries/pseudo-science/>
- Kageyama, Y. (2003). Openness to the unknown: The role of falsifiability in search of better knowledge. *Philosophy of the Social Sciences*, 33(1), 100–121. <https://doi.org/10.1177/0048393102250307>.
- Miller, D. (2006). *Out of error: Further essays on critical rationalism*. Ashgate.

- Popper, K. S. (1959). *The logic of scientific discovery*. Hutchinson & Co.
- Ruse, M. (2018). Creationism. In E. Zalta (Ed.), *The Stanford encyclopedia of philosophy* (Winter ed.). Retrieved November 2020, from <http://plato.stanford.edu/archives/win2018/entries/creationism/>
- Thornton, S. (2019). Karl Popper. In E. Zalta (Ed.), *The Stanford encyclopedia of philosophy* (Winter ed.). Retrieved November 2020, from <http://plato.stanford.edu/archives/win2019/entries/popper/>

FORMAL SCIENCES

Formal sciences are disciplines based on formal systems. These include logic, mathematics, and the theoretical fields of information theory, statistics, and computer science. Also included in the formal sciences are such areas as game theory, decision theory, and some areas of linguistics. Formal sciences use *a priori*, as opposed to empirical, methodology. Major advances in formal science have often led to major advances in the empirical sciences. The formal sciences are essential in the formation of hypotheses, theories, and laws and help us discover how things work, how people think and act.

Formal sciences are often considered the third branch of science, alongside natural science and social science. They came into existence long before the formulation of scientific method, since we know of mathematical texts dating back to ancient times. These include texts relating to Babylonian mathematics from around 1800 B.C.E., from Egyptian mathematics dating from around 1600 B.C.E., and examples of Indian mathematics dating back to around 1000 B.C.E. Other cultures have added their input along the way, including the Greeks and Islamists, as well as the Chinese and Japanese, who worked independently from one another.

Hundreds of years later, a branch of formal science called *logic* came into being. Like mathematics, logic originated from multiple cultures and methods. The first record of logic arose in India in the 6th century B.C.E., followed by China in the 5th century B.C.E. Modern logic originated from Greece and then Islam. It was during the 17th century C.E. that mathematics reached its zenith, with the earliest

studies of probability theory. Formal science in the modern world is now used in engineering, computer science, and many other disciplines.

This entry describes how the formal sciences developed, starting with the backbone of all the formal sciences, mathematics, and its various branches. Then the foundations of logic as a formal science, and its development, are examined and discussed the ways in which formal sciences are helping to shape the modern world.

Why Formal Sciences Differ From Other Sciences

One reason that mathematics, in particular, enjoys a special distinction above other sciences is because its laws are absolutely certain and incontrovertible. Those of other sciences are open to debate, and their theories in danger of being overthrown in favor of newly discovered facts.

As opposed to empirical sciences—the natural and social sciences—formal sciences do not involve themselves in empirical procedures. They do not presuppose knowledge of contingent facts or describe the real world. In this way, formal sciences are logically and methodologically *a priori*, their content and validity wholly independent of empirical procedures. Although formal sciences are conceptual systems, this does not place them apart from the real world, and they certainly have a relationship with it.

Formal statements can have a place in all possible, conceivable worlds; for example, there are a wide variety of mathematical axioms that categorize different incompatible worlds. This is in contrast to empirical theories, for example, the theory of relativity or evolution. These theories do not hold in all possible worlds, and in fact may not even turn out to hold in this world. It is this universal property that makes the formal sciences applicable to all domains and useful to all empirical sciences.

Formal sciences, being nonempirical in nature, are structured in terms of sets of axioms and definitions from which other statements (theorems) can be deduced. Put another way, theories in formal sciences contain no *synthetic* statements (those based on our sensory data and experiences), simply *analytic* statements, which are true by definition and are self-explanatory.

Foundations of Modern-Day Mathematics

Mathematics as a formal science is the cornerstone of the other formal sciences—indeed, they could not exist until mathematics had developed into a relatively advanced form of study. From the very early history of mathematics, mathematicians believed that the physical world was constructed along mathematical lines. Pythagoras (c. 570–490 B.C.E.) believed that everything was related to mathematics, and that numbers were the ultimate reality. Isaac Newton (1642–1727) thought that God used mathematics to design the world.

Mathematics is the science of structure, order, and relation that originated from the elemental practices of measuring, counting, and describing the shapes of objects. It is concerned with logical reasoning and quantitative calculation, and it has developed increasingly along the lines of abstraction and idealization.

Mathematics As the Basis of the Physical Sciences

Since the 17th century, mathematics has been an indispensable tool for the physical sciences and technology. Seventeenth-century mathematician Pierre de Fermat is often considered the founder of the modern theory of numbers. He also discovered the fundamental principle of analytic geometry and devised methods for finding tangents to curves along with their maximum and minimum points—a methodology which provided the backbone for differential calculus.

Another early mathematician, Christiaan Huygens (born 1629) started what we now think of as mathematics in its modern form by introducing us to the first probability theory. He also founded the wave theory of light, discovered the rings of Saturn, and contributed to the science of dynamics.

His influential work of 1654, *De Circulie Magnitudine Inventa*, helped to establish his growing reputation as an influential European mathematician. A later work of 1673 postulated, similarly to Fermat, a theory on the mathematics of curvatures. He also offered complete solutions to the problems of dynamics, having based his theory on a formula that measured the oscillation of a pendulum and the oscillation of a body about a static axis.

Algebraic Theory

Early 19th-century German mathematician Carl Friedrich Gauss was a pioneering mathematician both in his native Germany and abroad. His major mathematical publication was considered the first systematic textbook on algebraic number theory.

In *Disquisitiones Arithmeticae*, he gave an account of modular arithmetic, quadratic polynomials, and a theory of factorization. This book contained natural generalizations that led to the development of number theory, setting its agenda for much of the 19th century. He set forward a fundamental theorem of algebra, which stated that every polynomial equation with real or complex coefficients has as many roots as its degree (the highest power of the variable). He reached his conclusion mainly by critiquing earlier attempts to provide this proof.

Gauss later published further works on number theory, the mathematical theory of map construction, and other varied subjects. Gauss and French mathematician and astronomer Pierre-Simon Laplace (1749–1827) also worked together to develop the mathematical theory of statistics, which looked at statistics as it related to insurance and governmental accounting procedures. Mathematical statistics would not become a recognized mathematical discipline until the first part of the 20th century.

Mathematical Logic

The start of the 20th century was marked by a shift toward abstract thinking and generalization in mathematics, where notions of *self-evident* truths were largely discarded in favor of a move toward logical concepts such as *completeness* and *consistency*. The early part of the century saw the rise in mathematical logic which built on the advances made earlier by Gottlob Frege.

Austrian mathematician Kurt Gödel (1906–1978), however, soon endeavored to put severe constraints on what could and could not be solved in mathematics. His famous incompleteness theorems stated that in any reasonable mathematical system there will always be *true* statements that cannot be proved.

The Beginnings of Artificial Intelligence

Alan Turing (1912–1954) spent many years trying to clarify and simplify Gödel's proof, something which led him to conclusions which many saw as even more devastating than Gödel's. It conjectured that some purely mathematical questions are undecidable—that there is no one algorithm that can, infallibly, give a correct *yes* or *no* answer to each instance of a problem. He posited that there is no general method which can tell us whether a given formula U is provable in K .

A spin-off of this work led to the development of the first computers and considerations of concepts such as artificial intelligence (AI).

Game Theory

Following World War II, after the deliberate destruction of mathematics in Germany and Austria by the Nazis, the focus of mathematics moved to America and in particular to Princeton. Some of the brightest mathematical brains of the century could be found here, including John von Neumann and Albert Einstein. Von Neumann (1903–1957) made vast advances in quantum theory and played a significant role in the development of nuclear physics. He is also remembered for being an early pioneer of game theory.

Number Theory

Andre Weil (1906–1998) was another war refugee to Princeton who is noted for his theorems which allowed connections to be made between algebra, number theory, geometry, and topology. Alexander Grothendieck (1928–2014) was a protégé of Weil, a structuralist who became interested in the hidden structures within mathematics. He created a new language, which enabled mathematical structures to be seen in a new light, allowing for new solutions in number theory and geometry. This was termed the *theory of schemes*.

Paul Erdős was another inspirational 20th-century mathematician who worked on the problems of number theory, set theory, graph theory, and probability theory. The 20th century also saw the complex field of dynamics—an area first developed

by two French mathematicians, Pierre Fatou and Gaston Julia—receive significant attention.

Beginning of Mathematical Sciences

By the mid-20th century, mathematics had broadened its horizons to encompass the rising class of mathematical sciences and engineering disciplines, including systems engineering and operations research. These sciences relied on research in electrical engineering and benefited from the emergence of electrical computing. This in turn brought about information theory, numerical analysis (also known as *scientific computing*), and theoretical computer science. Theoretical computer science was also semi-reliant on mathematical logic and the theory of computation. These fields have seen mathematics become a major profession involving thousands of jobs.

Statistics and Probability

Pierre-Simon Laplace, mentioned earlier, was a French mathematician and astronomer whose work was pivotal to the development of mathematical astronomy. His early work in the 1770s on differential equations and work on the philosophical concepts of probability and statistics were groundbreaking. His paper "*Memoire Sur la Probabilité des cause par les Événements*" was published in 1774, in which Laplace indicated that probability was crucial to repairing defects in knowledge.

Laplace regarded analysis as a means for attacking physical problems, a belief that spurred him on to develop the *necessary analysis* theory. For Laplace, the results of an argument could be considered true without recourse to the steps which led to this conclusion. Solutions arrived at by any means would do, so long as they were true.

The German mathematician and philosopher Gottlob Frege's 1812 publication *Theorie analytique des Probabilités* was an important milestone in statistical theory. And in 1819, he published an account of his work on probability. This involved devising a method of estimating the ratio of the number of favorable cases compared to the whole

number of possible cases. His theory involved treating the successive values of any function as the coefficients in the expansion of another function, in reference to an alternative variable. This creates a probability-generating function.

The methods devised by Laplace, working in conjunction with Gauss, led to a normal probability distribution called the *Laplace–Gauss distribution*. Probability theory was in fact to become one of his life’s endeavors as illustrated in one of his well-known formulas, the *rule of succession*, where one trial has two possible outcomes: *success* and *failure*. If we assume that little or nothing is known a priori about the relative possible outcomes, Laplace’s formula determined the probability that the next trial would be a success.

$$\text{Probability that the next outcome is a success} = \frac{s + 1}{n + 2},$$

where s is the number of previously observed successes and n is the total number of observed trials. Using this formula Laplace calculated that the probability of the sun rising the next morning, given that it had never failed to in the past, was

$$\text{Probability that the next outcome is a success} = \frac{d + 1}{d + 2},$$

where d is the number of times the sun has risen in the past. The result has been derided as absurd, with some authors concluding that the rule has no meaningful use. Laplace’s other developments in pure and applied mathematics included discussion of the theory of determinants and proof that every equation of an even degree must have at least one quadratic factor. He also performed notable achievements in the areas of fractions, probabilities, and definite integrals.

Algorithms and Data Structures

Formal algorithms have existed for millennia, for example, Euclid’s algorithm for finding the greatest common divisor of two numbers, an algorithm that is still used today in computation. It was not until 1935, however, that Alan Turing, Stephen Kleene, and Alonzo Church formulated the term *algorithm* for computational purposes.

Algorithms are computable steps required to achieve a desired result and are particularly suited to computers because they can be computationally processed. Algorithms often rely on external information that is transformed into data and arranged into *data structures*. Computational theory, which started with Alan Turing, regards some problems as beyond the ability of a machine to process efficiently, which means that computer science needs to deal heavily in math, logic, and philosophy.

Information theory was included as a new discipline in 1948 with a mathematical theory of communication put forward by Claude Shannon (1916–2001) in his paper “A Mathematical Theory of Communication,” which proposed binary digits for coding information.

Shannon stated that all information has a *source rate* that can be measured in bits per second. Information theory is a communication theory devoted to the problems of coding. It is unique in its use of numerical measure of the amount of information gained when the content of a message is learned. The theory relies heavily on the mathematical science of probability and is often applied to other probabilistic studies, including random noise and signal detection.

Emergence of Neuroscience

The Canadian psychologist Donald Hebb (1904–1985) introduced the mathematical model of learning in the brain, which was an attempt to understand the function of neurons contributing to psychological processes. His *Organization of Behavior* is considered his most significant contribution to the world of neuroscience and is where he connected the biological function of the brain as an organ together with the higher function of the mind. This was based on earlier research done by Hans Berger, who had cast doubt on the Pavlovian model of perception on response after new evidence appeared to show that the brain exhibited activity even when there was little stimulus. *Hebbian* learning is said to follow the following model:

If an axon of a cell (A) is close enough to excite cell B, and is repeatedly or persistently firing it, a metabolic change will be actioned in one or both

cells. This means that A's efficiency as one of the cells firing B will increase.

In other words, neurons that *fire together wire together*, commonly referred to as *Hebb's law*. The combination of neurons making up a processing unit are termed *cell assemblies*, their connections together making up the ever-changing algorithms that dictate the response of the brain to stimuli. This not only opened up a model for the working of the mind, in terms of the processing of stimuli, but also paved the way for the creation of computational machines that could mimic the biological processes of a living nervous system. Despite that fact that this form of synaptic transmission was later found to be chemical, modern artificial neural networks are still today based on the transmission of signals via electrical impulses.

Following the work of Hebb came the fields of neural networks and parallel distributed processing. In the early 1970s, Stephen Cook and Leonid Levin proved the existence of practically complete *nondeterministic polynomial time* problems, a revelation in computational complexity theory.

Foundations of Logic

Along with mathematics, formal logic is one of the oldest sciences in the realm of formal sciences. It was used explicitly to analyze methods of reasoning in India from 600 B.C.E., China in 500 B.C.E., and Greece between the 4th- and 1st-century B.C.E. Logic was given a sophisticated makeover by the Greeks and in particular by Aristotle in Aristotelian logic. Subsequently, logic was taken further by Islamic logicians and Indian philosophers.

The term *logic* originated from the Greek word *logos*, which can mean *sentence*, *discourse*, *reason*, or *rule*. Logic is the use and study of valid reasoning and features prominently in the subjects of mathematics, philosophy, and computer science.

First established as a formal discipline by Aristotle, the study of logic was one of the classical trivium, which also included rhetoric and grammar. Aristotle's system of syllogistic logic remained predominant in the West until the mid-19th century, when interest in the foundations of

mathematics led to the development of symbolic logic, now termed *mathematical logic*.

Aristotle's works on logic were later grouped under the name *Organon* or *instrument*. Aristotle, however, used the term to represent verbal reasoning.

Aristotelian logic posited that notions in isolation cannot express truth or falsehood. It is only when they are combined in a proposition or combination of words in rational speech that their truth or falsity can be determined. All Aristotelian logic revolves around a single notion: *deduction* (syllogism). A deduction equates to a modern notion of logical consequence: *X* must, of necessity, result from *Y* and *Z* if *X* cannot be false when *Y* and *Z* are true. This is also a definition of *valid argument*. Syllogisms are structures of sentences which can be meaningfully called *true* or *false*.

According to Aristotle, every such sentence must have the same structure and contain a subject and a predicate and must either affirm or deny a single predicate of a single subject. A classic example is his syllogism is

All men are mortal. Socrates is a man. Therefore Socrates is mortal.

So long as the premises are true, the conclusion will be too.

Aristotle's logic is a kind of *formal* logic, which as previously stated involves an inference that applies to a wholly abstract rule (i.e., a rule that is not about any particular thing or property). It involves arguments having the form of sentences of formal grammar set out in a logical language so that its content can be used in formal inference.

Because indicative sentences of ordinary language have a wide variety of form and complexity, Aristotle's formal logic requires us to ignore those features that are irrelevant to logic, replacing conjunctions such as *but* with logical conjunctions such as *and*. It also requires us to replace ambiguous expressions such as *any* with standard logical expressions such as *all*. There also needs to be recourse to schematic letters; for example, "all A's are B's" illustrates the logical form common to the sentences "all men are mortals," and so on.

Aristotle used variable letters to represent inferences in *Prior Analytics*.

Necessity in Logic

The principles of logic are noncontingent and do not depend on any particular accidental features of the world. Physics and other sciences investigate how the world is. Physicists may tell us that no signal can travel faster than the speed of light, but if the laws of physics were different this might not have been true. Similarly, biologists study how dolphins behave, but if the course of evolution had been different dolphins might not even have existed. Empirical science is contingent on how things are or could have been.

The principles of logic, on the other hand, depend only on reason, and their validity does not depend on any contingent features of the world. If p , then p is necessarily true. “If it is raining, then it is raining” is true. This is true even if it isn’t actually raining. Even if the laws of physics change, this is still true.

Formal logic is concerned with formal systems of logic especially constructed for carrying out proofs where language and rules are clearly defined. *Sentential logic* (also known as *propositional logic*) and *predicate logic* are both examples of formal systems of logic.

There are many reasons for studying formal logic: First, it helps us identify patterns of good reasoning and patterns of bad reasoning. Second, it can help with critical thinking. Formal systems of logic are also used by linguists to study natural languages.

Computer scientists employ formal systems of logic in research relating to AI. Finally, philosophers like to use formal logic when dealing with difficult philosophical problems, in order to make their reasoning explicit and precise.

Deductive Reasoning and Inductive Reasoning

Deductive reasoning is concerned with what necessarily follows from given premises (if a then b). *Inductive* reasoning is the process of deriving a reliable generalization from observations and is sometimes included in the study of logic. *Validity* is another important point: An inference is deductively valid *if and only if* no possible situation exists in which the premises are true but the conclusion is false.

Inductive arguments, however, are neither valid nor invalid, since their premises allow only for a probability of certainty, without concluding anything. While the notion of deductive validity can be stated for systems of logic in terms of semantics, inductive validity requires sets of observations, some less formal than others. These can use mathematical models of probability.

Logical systems must be consistent; theorems within a system cannot contradict one another. They must also be valid in that the system’s rules of proof never allow for a false inference from the premises. Logical systems must have the property of *soundness*, which is only possible if the inference rules prove only formulas that are valid semantically. In general, this means having the property of *truth*—although this is not always the case.

An argument can be said to be sound only if the argument is valid and all of the premises are true. For example, “All men are mortal,” “Socrates is a man,” and “Therefore Socrates is mortal.” This is a valid argument because the conclusion is true, based on the premises; in other words, the conclusion follows from the premises, and since the premises are in fact true, the argument is sound.

A logical system must be *complete*, meaning that if the formula is true, it can be proven (if it is true, it is a theorem of the system). Although most logical systems possess these properties, not all of them do.

Jan Łukasiewicz, a Polish logician, stated that the introduction of variables was one of Aristotle’s greatest inventions. According to Aristotelian logic, only the logical principles stated in schematic terms belong to logic—not the concrete examples. For instance, the concrete terms *man* and *mortal* are analogous to the substitution values of the placeholders A, B, C—termed the *matter* of the inference.

Along with the Aristotelean syllogism, modern symbolic logic is an example of formal logic. Symbolic logic is the study of symbolic abstractions that capture the formal features of logical inference. It is often divided into two branches, *propositional logic* and *predicate logic*. Mathematical logic is an extension of symbolic logic and particularly applies to the study of model theory, set theory, proof theory, and recursion theory.

In 1854, the English mathematician George Boole published *An Investigation of the Laws of*

Thought on Which are Founded the Mathematical Theories of Logic and Possibilities, which introduced the principles of symbolic logic and what is now referred to as *Boolean logic*. Boolean logic is a complete system for logical operations.

Today it has many applications in computer technology and forms the basis of digital electronics. It involves a system of algebra in which all values are reduced to either true or false, and it is these true/false rules that govern the foundation of all electronic circuits in a computer. Just as *add*, *subtract*, *multiply*, and *divide* are the primary operations of arithmetic, so *and*, *or*, and *not* are the primary operations of Boolean logic.

Often considered the founder of modern logic and of analytic philosophy, Gottlob Frege, mentioned earlier in this entry, found the language of Boole unsuitable for being too unambiguous. For Frege, Boole's logic, which was divided into primary and secondary logic, was unable to deal adequately with multiple generalities. Frege found the method of analyzing propositions in terms of subjects and predicate concepts imprecise and antiquated. He replaced subject/predicate with the notions of *function* and *argument* and defined a *concept* as a function that has a truth value, with true or false as its value for any object as argument. His logic was based on axioms, beginning with a fixed number of axioms from which all other logical truths could be derived.

In 1869, Frege published *Begriffsschrift* (*concept-script*), which put forth quantifier notation. This predated modern predicate logic by dealing with quantifiers using variable bindings. Frege attempted to show that all basic arithmetic truths can be derived from purely logical axioms. In his *Basic Laws of Arithmetic, volume I*, he presented a new logical language. In his paper "A Logical Inquiry," first published in 1918, Frege saw logic as the means to discover the laws of *truth* itself, whereas discovering different truths is the task of all the sciences.

Frege was to play a crucial role in the development of modern logic and analytic philosophy. His work in logic is often seen as the representation of a fundamental break between contemporary approaches and the older Aristotelian tradition. He invented modern quantificational logic and created the first fully axiomatic system for logic that was complete in its treatment of propositional

and first-order logic. When it came to mathematical logic, he was a proponent of the thesis that mathematical truths are logical truths. His distinction between sense and reference of linguistic expressions was groundbreaking in the work of semantics and language, profoundly influencing thinkers such as Bertrand Russell, Rudolf Carnap, and Ludwig Wittgenstein.

Rival Conceptions of Logic

Logic arose from efforts to achieve correctness in argumentation. Some logicians, however, believe that logic should only be concerned with the right forms of inference. The American philosopher Thomas Hofweber believes that logic should not cover good reasoning as a whole; that should be the job of rationality. It should instead deal with inferences whose validity can be traced back to formal features of the representations involved in inference, whether mental, linguistic, or otherwise.

Informal logic looks at natural language arguments and the study of fallacies, Plato's dialogues being good examples. Informal logic is used to mean the same thing as critical thinking and sometimes refer to the study of reasoning and fallacies in everyday life.

The Importance of the Formal Sciences

Formal sciences are essential to the sciences in general. Mathematics in particular plays a crucial role in the expression of scientific models. Observing and collecting measurements, as well as predicting and hypothesizing, usually require extensive use of mathematics. Arithmetic, geometry, algebra, trigonometry, and calculus are all, for example, essential to physics.

Virtually all branches of mathematics have applications in science, including in so-called pure areas of number theory and topology. Statistical methods, which comprise mathematical techniques for summarizing and analyzing data, enable scientists to assess levels of reliability and range of variation in experiments. Statistical analysis also plays a major role in the natural and social sciences.

The formal computational sciences enable us to gain a better understanding of scientific problems, better even than mathematics alone. In fact, according to the Society for Industrial and Applied

Mathematics, computation is nowadays as important as experiment and theory for advancing our knowledge of science. The theorems and formulas obtained by logical means that presume axiomatic systems, rather than the combination of empirical observation and logical reasoning, has become known as the *scientific method*.

Major advances in formal science have often led to major advances in the empirical sciences. The formal sciences are essential in the formation of hypotheses, theories, and laws and help us to discover how things work, how people think and act.

Thora L. Fitzpatrick

See also Computer Science; Information Theory; Logical Theory; Mathematics, 19th Century; Mathematics, 20th Century; Mathematics, Enlightenment

Further Readings

- Churchman, C. W. (1940). *Elements of logic and formal science*. J.B. Lippincott.
- Crease, R. P. (2011). *World in the balance: The historic quest for an absolute system of measurement*. W. W. Norton.
- Davis, M., Sigal, R., & Weyuker, E. J. (1994). *Computability, complexity, and languages: fundamentals of theoretical computer science* (2nd ed.). Academic Press.
- Fara, P. (2009). *Science: a four-thousand year history*. Oxford University Press.
- Gleick, J. (2010). *The information: A history, a theory, a flood*. Pantheon.
- Halmos, P., & Givant, S. (1998). *Logic as algebra*. MAA Press.
- Hamilton, A. G. (1978). *Logic for mathematicians*. Cambridge University Press.
- Kuhn, T. (1970). *The structure of scientific revolutions*. University of Chicago Press.
- Shannon, C., E. & Weaver, W. (1964). *The mathematical theory of communication*. University of Illinois Press.
- Suppes, P. (1957). *Introduction to logic*. Van Nostrand.
- Quine, W. V. O. (1953). Two dogmas of empiricism. In *From a logical point of view*. Harvard University Press.
- Woodward, J., & Ross, R. (2021). Scientific explanation. In E. Zalta (Ed.), *The Stanford encyclopedia of philosophy* (Summer ed.). <https://plato.stanford.edu/entries/scientific-explanation/>
- Ziman, J. (1978). *Reliable knowledge: An exploration of the grounds for belief in science*. Cambridge University Press.

FRAMEWORK

A framework is an abstract logical structure of meaning that guides the development of a study enabling a researcher to link their findings to a body of knowledge. A *theoretical framework* concerns a framework of study that is based on a theory. A *conceptual framework* is a term used in a study that has its roots in a specified conceptual model.

A conceptual framework requires that a researcher uses concepts from the available literature to establish evidence to support the need for a certain research question. If these links have already been established through valid research, this ready-made theoretical framework can be used as a map to help guide other scientists with their own research questions.

Theoretical frameworks provide a perspective through which it is possible to examine a topic. There are no right or wrong theoretical frameworks, since a topic can be looked at from multiple perspectives. Frameworks are there to guide research, determine which things should be measured, and define what kinds of statistical relationships one should be alert for.

Theoretical frameworks are critical in deductive, theory-testing types of study. They consist of interrelated concepts, together with their definitions, along with existing theories that are used for a particular study. Frameworks need to express an understanding of the theories and topics relevant to the topic of research and relate them to broader fields of knowledge.

Frameworks are not *off the peg* or readily available in the literature, since they must be applied to the case at hand.

A theoretical framework strengthens a research study in the following ways:

- It provides an explicit statement of theoretical assumptions which permit the reader to critically evaluate them.
- It connects the researcher to the existing knowledge and guides them with a relevant theory. This provides a basis for the hypothesis and choice of research methods.
- It articulates the theoretical assumptions of research study, forcing researchers to address *why* and *how* questions. Theoretical frameworks allow for us to move from simply *describing*

observed situations or phenomena to generalizing various aspects of a phenomenon or occurrence.

- Having a structure helps to apply limits to those generalizations. A theoretical framework identifies the key variables that influence a phenomenon, encouraging examination of how those variables might differ and under what circumstances.
- Frameworks are useful organizing devices in empirical research. A conceptual framework can be applied to deductive, empirical research at micro- or individual-level study.

Conceptual Frameworks

Conceptual frameworks can be thought of as abstract representations connected to the goal of the research project. Patricia Shields and Nandhini Rangarajan use American football play as a good metaphor for a conceptual framework. They define a conceptual framework as the way ideas are organized to achieve an objective. Football play, like a conceptual framework, is of course connected to a purpose. In American football, a play is a strategy or on-the-ground plan of action used to move the ball down the field. A *play* occurs at kickoff or a *snap* during a down, and plays can be basic or complex. The components of a play are subdivided into many smaller parts, just as the parts of a framework should be divided into smaller aspects.

According to Shields and Rangarajan, empirical inquiry is akin to coaching football and inventing plays. Researchers should see themselves as coaches who need to develop frameworks (plays) as they work through their literature and define research objectives. Shields and Rangarajan introduce five main types of conceptual frameworks designed to bring coherence to all aspects of the research process. Each conceptual framework is defined to match a specific research purpose. The five frameworks are *working hypotheses*, *categories*, *practical ideal type*, *models of operations research*, and *formal hypotheses*.

Working Hypotheses: Used for Exploratory Research

A working hypothesis is a hypothesis that has been provisionally accepted as a basis for further research. In using a working hypothesis, it is hoped that even if the hypothesis fails, a tenable

theory may be produced. Like all hypotheses, a working hypothesis is constructed in terms of explanations that can be linked to the explanatory research. Charles Sanders Peirce (1839–1914) set forth his belief that an explanatory hypothesis is useful not only for the tentative conclusion it entails but also as a starting point for further research. This idea of justifying a hypothesis as worthy at the research level, not just for its use as a logical conclusion, is central to the notion of a working hypothesis. This notion was later elaborated by pragmatist John Dewey (1859–1952).

Categories: Used for Descriptive Research

Descriptive research relates to the characteristics of a phenomenon under study. It does not answer “why?” “when?” or “how?” the characteristics have occurred but rather answers the “what?” question; namely, what are the characteristics of the phenomenon being studied? These characteristics generally make up a kind of categorical scheme also known as *descriptive categories*. An example we could use is that of the periodic table, which categorizes the elements. Scientists have used their knowledge of the characteristics of protons, electrons, and neutrons to create this categorical scheme. We now take this scheme (the periodic table) for granted; however, it was originally based on descriptive research. Descriptive research tends to precede explanatory research—as illustrated in the fact that scientists use the descriptive elements of the periodic table to explain chemical reactions and make predictions about what will happen when elements are combined.

Practical Ideal Type: Used for Gauging

A practical ideal conceptual framework involves that there be a set of criteria suitable for judging a program process or activity. These criteria can help to determine best practices. The term *practical* relates to the idea type because the model is subject to revision.

Models of Operations Research: For Use in Decision-Making

Mathematical models and equations and computer models are often used to describe the past behavior of a process and predict the future path

of a process within the boundaries of the model. In certain cases, the probability of an outcome rather than its specific outcome will be predicted, as in much of quantum physics. The goal of operations research is to provide a framework for constructing models of decision-making problems, find the best solutions, and implement those solutions in an effort to solve the problems. The *problem* is the occurrence or situation that warrants enquiry. The decision maker is the individual or group responsible for making decisions regarding its solution. The problem may be abstract or real and may involve current operations or proposed operations.

Formal Hypotheses: For Use in Explaining and Predicting

A formal hypothesis consists of concepts that can be connected, and their relationship tested. Numerous hypotheses amalgamate to form a conceptual framework. When sufficient data and evidence are gathered to support a hypothesis, it becomes a working hypothesis. Proposing a formal hypothesis is an extremely important part of the research project, since the thesis or study is at the mercy of the hypothesis.

Explanation is the most common type of research purpose in empirical research. The formal hypothesis is the framework associated with explanation. Explanatory research generally focuses on *why* a phenomenon has been caused to happen. Formal hypotheses provide possible explanations in answer to the questions that are tested by collecting data and then assessing the evidence, which is usually quantitative via statistical testing. An example of this could be where a researcher wishes to determine the factors contributing to house fires in American cities. Three factors may be put forward as possible influences on residential fires: population, building makeup, and the environment. These factors would become the conceptual framework they used to achieve the research goals to explain the influences of residential fires in U.S. cities.

A *prediction* is a forecast or statement about the way things will happen in the future, usually, although not always, based on experience or knowledge. A prediction is more specific than a forecast since it makes a statement about what

may happen or what is expected to happen, while a forecast may cover a range of possible outcomes. Predictions are useful when making plans about possible developments. More formal predictions can be made by testing formal hypotheses using statistical methods. Formal hypotheses are created by connecting systematic knowledge of an area giving clear reasons why the predicted relationship might exist. Predictive inference is part of statistical inference. Established science makes useful predictions that can be highly accurate. New theories allow for predictions which can be disproved if they are not borne out. Scientific predictions are rigorous and generally quantitative statements forecasting what will happen under certain conditions. For example, if an apple falls from a tree it will be pulled toward the center of the earth by gravity and will do so according to the specified and constant acceleration. Scientific method is built on testing statements that are logical consequences of scientific theories. This is done through observation and experiment.

These conceptual frameworks, as put forward by Shield and Rangarajan, are not exhaustive, as they themselves point out. Such frameworks may not be applicable for inductive forms of empirical research, either. They merely provide a useful point of departure for those new to research when setting out their own research design.

Frameworks Provide a Lens

Frameworks are important in explanatory studies where about the areas to be researched are unknown beforehand. Frameworks are especially helpful in these instances because, even if not much is known about a topic and no matter how unbiased the position we *feel* we are taking, it is difficult not to have preconceived notions—even a very general preconceived notion about the subject we are about to tackle. If, for example, a researcher has a fundamental view about how people in general behave (perhaps they has an inclination to think they are lazy and unhelpful), this may cause bias when researching a topic that involves human nature. This researcher may fail to notice things that don't fit in with this preconceived framework. By making a framework explicit at the start, other frameworks can be tried

out to see the organizational situation through a different lens.

Isaiah Berlin (1909–1997) believed that philosophers and authors view the world either from a fox’s point of view or from a hedgehog’s—two types of animals he used to describe the different ways in which we see the world. They are a way of describing the analytical tools, or conceptual frameworks, we each use to make conceptual distinctions and organize ideas. According to Berlin, the philosopher Plato, the playwright Henrik Ibsen, and the novelist Fyodor Dostoevsky were *hedgehogs*, or types of people who view the world in terms of a single idea. Foxes, on the other hand, according to Berlin, incorporate a pluralistic view of the world, in that they draw on a variety of experiences. Rather than pinning things down to a single idea, they see it through multiple, sometimes conflicting lenses. He regarded Aristotle, Shakespeare, and the novelist James Joyce as being in this category. Berlin took the title of his essay “The Hedgehog and the Fox” from a reference to an ancient poetical work by Archilochus, who describes the fox as knowing many little things, while the hedgehog knows just one big thing. Our individual conceptual framework is based on our beliefs, experiences, and knowledge. A theoretical framework helps researchers to understand and overcome their personal biases so they can gain better insight into their results. They can also learn to use other frameworks during the course of their study; for example, they could learn to stage their research from more of a fox’s or more of a hedgehog’s type of view.

Types of Research

Frameworks are useful in both applied and basic research, applied research being designed to solve a particular problem in a particular circumstance. Applied research refers to scientific study and research that seeks to solve a practical problem. This could include efforts to find solutions to everyday problems and develop new technologies. Applied research generally employs empirical methodologies.

Frameworks are also important for basic or *pure* research, which is designed to find the principles that underlie certain situations. Basic

research involves discussion about ideals, programs, and methodologies.

Frameworks have also been used to good effect in conflict theory, allowing for a way to determine the balance required to reach a resolution. Within conflict frameworks, both visible and invisible variables are at play, with both parties having a subjective viewpoint. Adopting frameworks allows us to focus on specific variables and define an objective framework that the research will take in analyzing and interpreting the data to be gathered. It allows us to build knowledge by validating or challenging theoretical assumptions, thus enabling us to take a practical approach to conflict resolution, for example. Mediation is built upon a commitment to developing conflict resolution skills both individually and collectively, in order to develop equitable relationships. The development of these skills requires an acceptance of plurality or an agreement that there are a variety of possible ways to determine both the process and outcome of dispute resolution. That there are potentially a number of valid points of view poses a fundamental challenge at the philosophical level to established assumptions and increases the likelihood of an agreement.

How to Develop a Framework

First, we need to examine the research problem, since this anchors the entire study and forms the basis of the framework.

Next is to brainstorm the key variables in the research and answer the question: What factors contribute to the presumed effect?

A review of related literature should then be undertaken to find answers to the research question.

The next step is to list the constructs and variables that might impact the study and group them into independent and dependent categories.

Finally, a researcher should review key theories and choose the theory or theories that best explain the relationships between the key variables. The final step is to discuss the assumptions used to limit the scope of the relevant data.

Thora L. Fitzpatrick

See also Concept; Conceptual Analysis; Explanation; Generalization; Worldview

Further Readings

- Berlin, I. (1953). *The hedgehog and the fox*. Simon & Schuster.
- Botha, M. E. (1989). Theory development in perspective: The role of conceptual frameworks and models in theory development. *Journal of Advanced Nursing*, 14(1), 49–55.
- Dewey, J. (1938). *Logic: The theory of inquiry*. Holt, Rinehart and Winston.
- Kaplan, A. (1964). *The conduct of inquiry: Methodology for behavioral science*. Chandler.
- Peirce, C. S. (1878). How to make our ideas clear. *Popular Science Monthly*, 12, 286–302. (Reprinted widely, including *The essential Peirce*, Vol. 1, pp. 109–123)
- Shields, P., & Rangarajan, N. (2013). *A playbook for research methods: Integrating conceptual frameworks and project management*. New Forums Press.
- Shield, P., & Tajalli, H. (2006). Intermediate theory: the missing link in successful student scholarship. *Journal of Public Affairs Education*, 12(3), 313–334. <https://doi.org/10.1080/15236803.2006.12001438>
- Shields, P. (1998). Pragmatism as a philosophy of science: A tool for public administration jump up. In J. D. White (Ed.), *Research in public administration* (pp. 195–225). Emerald Group.

G

GAME THEORY

Game theory is a mathematical system of reasoning, referring to cases where two or more decision-making agents with a set of preferences associated with the various possible outcomes participate in a process. The actions of each intervenient are critically affected by the decisions made by the other/s, and the final result depends on the totality of strategies adopted. Each game is characterized by a set of information available to participants, a predetermined number of rules that limit the strategic choices of decision makers, and a range of possible outcomes or *payoffs*. In broad terms, it may be said that game theory is based on the notion that people's choices are a result of well-defined preferences and an analysis of the relationship between their own choices and the decisions of others. Given this profile, game theory is a tool of high analytical potential regarding situations of conflict and cooperation based on strategic behavior, specifically referring the ways how people interact and decide. By studying the strategic behavior of agents, it is possible to sketch answers to questions regarding what is decisive for the existence of cooperation between the parties or when it is rationally more beneficial to compete.

The Theory

Generically, game theory models operate on the basis of concepts related to *rational choice theory*, with decision makers taking the course of action

considered the most suitable according to a certain preference scale and choices being made without any fundamental qualitative restrictions. Various types of strategies are identifiable. The *dominant strategy*, so called, occurs when a course of action is possible, better than any other for one single player, regardless of the actions of others. A *pure strategy* exists when it is possible to outline the moves that a player will make when facing some situation. A *mixed strategy* corresponds to assigning one probability to each possible pure strategy in the game. The variety of strategy named *max-min* posits players adopting a cautious stance and seeking to obtain the least negative of a group of poor results. The *tit-for-tat* strategy refers to a replicative-retaliatory attitude by players: Initially, the player cooperates, analyzes the behavior of others, and subsequently tends to reciprocate their actions. In the *grim-trigger strategy*, the player also cooperates the first time and insofar as the remaining players continue to collaborate, but if one of the others fails to cooperate, this is deemed sufficient cause for giving up and abandoning cooperation. Other noteworthy strategies are *collusion*, when there is an agreement between certain elements in order to deceive others and gain advantages in the game; *backwards induction*, when, starting from the end of the game, the player considers the different options and outcomes in order to determine the optimal sequence of actions; *forward induction*, when it is assumed that past strategies have been chosen rationally by several players and, following its analysis, it typifies the behavior of each one with a view to predicting

their future actions; and *Markov strategy*, when players proceed based only on descriptions that summarize the history of the game.

One classical problem of game theory, referred to as the *prisoner's dilemma*, remains a key analytical tool, capable of displaying how two individuals may rationally decide not to cooperate, even though circumstances presumably would tend to make cooperation advantageous for both. Initially suggested by Merrill Flood and Melvin Dresher in 1950, this dilemma was later formalized by Albert W. Tucker, who named the device exemplifying with the theme of imprisonment. Two prisoners accused of the same crime are kept separate, and each is given the option to confess. The authorities lack sufficient evidence to accuse them and hence are dependent on their testimonies. If neither one confesses, both will be given a sentence of 2 years. If both confess, they incur a sentence of 6 years. If only one of them confesses, he or she who does is released while the other will have an extended sentence of 10 years. Considering this situation closer we verify that, although each individual prisoner has advantage to confess (penalty of 0 or 6 years for confession, versus respectively 2 or 10 for nonconfession), joint nonconfession is the best option for the pair of players: penalty of 2 years each, instead of 6 each for joint confession. Notice that the best option is unstable, given the uncertainty for each party regarding the cooperativeness of the other. By contrast, a situation of mutual noncooperation corresponds to the so-called *Nash equilibrium*: No player can improve his or her outcome by unilaterally changing strategy, in spite of the existence of a better solution for both parties. However, it is not obtainable by simply acting otherwise because the improved results crucially depend also on changes in the behavior of other agents. Therefore, the problem basically consists in how to avoid falling into situations of generalized noncooperation, which subsequently become almost impossible to leave behind.

History

Some of the core ideas of game theory are identifiable as early as the 18th and early 19th centuries, in the works of James Waldegrave (1684–1741) and Augustin Cournot (1801–1877), but the

decisive developments occurred during the first two decades of the 20th century, with the works of Ernst Zermelo (1871–1953), Émile Borel (1871–1956), and John von Neumann (1903–1957). Von Neumann's first publication concerning games, in 1928, sought to demonstrate that in finite zero-sum games—that is, when the gain of one player corresponds to the loss of another—the solution for mixed strategies exists and may be obtained by applying mathematical techniques. In 1937, von Neumann presented a new model, this time based on Luitzen Egbertus Jan Brouwer's *fixed-point theorem* (presented in 1910); and in 1944, with *Theory of Games and Economic Behavior* by John von Neumann and Oskar Morgenstern, game theory acquired the status of a field of studies, extending its scope of influence to applied mathematics and economics. The mathematician John Nash (b. 1928) published in 1950 a series of articles in which he defined the concepts that became known as *Nash equilibrium* and the *bargaining problem*, giving renewed momentum to game theory, whose models started to be used in political science, economics, and psychology. Nash distinguished two types of games: cooperative ones, in which decision makers can combine in order to achieve a more favorable solution, and noncooperative games where the players' personal incentives prevail, leading to a less-than-excellent result that still configures a stable situation. According to the theorem of Nash equilibrium, participants seek a set of alternatives whereby everyone may be benefited by his or her strategy. Taking the competition as assumption, von Neumann concentrated its approach in zero-sum games, positing that even matches involving several people are analytically reducible to multiple sets of two. In turn, Nash chose to introduce the cooperative element, considering that through cooperation, individual gains may be maximized, and therefore, in the strategic analysis of the game, participants must take into account both the individual and collective perspectives.

Game theory has been applied extensively since the 1940s. Initially, it was used essentially for military purposes, notably in World War II, the Cold War, and the Korean War, application in other domains quickly followed, including physics, zoology, and social sciences. Indeed, Richard Dawkins (b. 1941) has importantly shown that

game theory is usable in studying some of the behavior patterns of genes in the evolution of species: In some cases, genes that cooperate evolve, obtaining higher individual gain. Since the 1970s, game theory is also used as an instrument of evolutionary biology, and its theoretical methods became very important in microeconomic theory as well, with the Nobel Prize in 1994 being awarded to three game theorists: John C. Harsanyi (1920–2000), who focused on situations where some decision makers have inside information as to important aspects of the game, corresponding to the so-called model of incomplete information; John F. Nash and Reinhard Selten (b. 1930), who in 1965 provided more in-depth knowledge of equilibrium, with the introduction of the concept of *subgame perfect equilibrium*. Nash identified three modes of application of game theory: economics, aiming at choosing one option based on the calculations of probability and maximization of utility; psychological, which by highlighting notions of incomplete information and expected utility emphasize the critical nature of intentions and expectations of the various participants; and finally the sociological mode, understanding the decisions taken by analyzing previous situations. Game theory application to economics should, however, be added with an important proviso: Its model of reasoning assumes a *strategic rationality*, whereby the conduct of each agent partially shapes the environment faced the others (and reciprocally), whereas economics' model of perfect-competition markets instead posits a simpler *parametric rationality*, each agent's conduct fundamentally irrelevant to the design of the environment faced by all of them.

Three successive generations are also identifiable in the approach to game theory: von Neumann and Morgenstern laid the foundations of the theory that aimed to be a tool for rationality analysis, referring to contexts in which the actions of the various decision makers are conditioned by expectations of possible reactions. Starting from the idea that policymakers adopt actions they rationally identify as more advantageous to themselves, the analysis of aspects such as the consequences of the available options, or the possible existence of alliances between players, was understood as a way of deduction of optimal strategies for each hypothesis. The second generation was led by Omar Khayyam Moore and Alan Ross

Anderson, who introduced one psychological component in the behavior of decision makers, leading to the concept of *subjective probabilities*: Players may not be fully informed, and so expected utility (both in terms of intent and of expectations) for each dimension plays a very important role. Robert Aumann (b. 1930) presented in 1987 what is considered the third approach, based on a combination of logical and subjective probabilities of players. According to Aumann, given that policy makers may adopt nonoptimal strategies, there is an undetermined number of possible solutions for each game, which implies introducing the notion of risk and calculating the probabilities of various outcomes. Aumann and Thomas Schelling (b. 1921) were awarded the 2005 Nobel Prize in Economics for their use of game theory in strategic analysis of international conflict resolution. Schelling pioneered the use of game theory with this purpose, including a group of studies on the nuclear arms race during the cold war. He presented concepts such as *conflict behavior*, *focal point*, and *credible commitment* that generically deal with various forms of subliminal and indeed mostly preconscious knowledge, but often very important regarding decision making within contexts of interaction. Credible commitment consists of a group of dispositions not strictly rational in character, yet still necessary both in order to prevent noncompliance in cooperative situations and also allowing participants to take advantage of the very plausibility that situations might evolve toward conflict, by benefiting from possible antagonists' risk-aversion.

Mention is also due to Robert Axelrod (b. 1943), who in *The Evolution of Cooperation*, published in 1984, highlighted the rule of *tit for tat*, or *pay in kind*, and posited a very efficient strategy in situations of prisoner's dilemma, containing an intuitive moral justification. His argument for the so-called evolutionary viability of *tit for tat* stresses that it verifies first of all the trait of being a very simple rule, one likely to emerge regardless of cultural components, altruistic feelings, and even the very existence of intentionality. This is indeed a decision-making algorithm originally presented by the mathematician Anatol Rapoport, implying to cooperate in principle, to retaliate whenever the interlocutor *misbehaves*—that is, doesn't cooperate—yet still to remain willing to

return to a cooperative attitude, once payback is consummated. Therefore, this is an overall *nice* (in principle cooperative) rule, *retaliatory*, yet still *forgiving* strategy. Tit for tat proved, in various tournaments of prisoner's dilemma, to be a highly successful option, besides favoring other cooperative strategies, which tended to be subject to predation by noncooperative, or *free-riding* strategies, and also inhibiting these by implying the limitation of their success: The typical tit-for-tat player is virtually impossible to deceive, which results from the defining algorithm's simplicity.

Continuing Developments

Robert Axelrod himself extracted a group of conclusions regarding the manifold and multiform implications of the reasoning just described. An important field of application is biology, where the actual fact of cooperation is usually recognized, in spite of the global frame of reasoning mostly suggesting a Darwin-inspired generalized *struggle for life*. The central question therefore concerns the very possibility of cooperation, and this theorizing is important precisely because it stresses simplicity, and so also the plausibility of emergence and subsequent viability of cooperative behaviors—assumed as unintended but a likely result of random mutations. Axelrod stresses that this variety of behavior may initially occur without any centralized decision or command order regardless of cultural devices such as moral norms, with no need for benevolence or mutual sympathy, and bottom line without any real intention of producing such cooperative result. However, he adds, those cultural devices, and psychological inclinations for *altruistic* or *benevolent* behavior indeed tend to emerge: more as a consequence of cooperation than as cause for it, although no doubt such devices do reinforce the tendency for a subsequent cooperative conduct.

In his 1997 work *The Complexity of Cooperation*, Axelrod returns to these themes, namely, assuming that the models used, the so-called *agent-based models*, which may be said to correspond to the generalization of problems typical of the *prisoner's dilemma* to n participants, have neither strictly inductive nor rigorously deductive characteristics, rather being mostly imagination-auxiliaries. Another noteworthy trait is precisely

the explicit recognition, by Axelrod, of the fact that his work had been discussed with actual policymakers, namely, people making decisions relevant for the United States's foreign policy: This interaction was not only a source of enrichment for Axelrod's assumptions but also of his theorizing's becoming susceptible of influencing the way agents proceed. In short, this is explicitly not a theory regarding *outside* reality and merely *about* it, but rather a theory whose components are quickly, directly reabsorbed by the realities studied and coming to integrate them. This induces the problem of knowing to what extent game theory itself may be, and indeed is, taken into consideration by strategic decision makers, and therefore, the question emerges of the theory's strictly *performative* validity: It may occur that it is only tenable insofar as agents become persuaded of its validity, thereafter proceeding in accordance with this belief.

Finally, a group of studies must be referred to, intimately related with the prisoner's dilemma model and also corresponding to the generalization of its reasoning scheme to situations where not just two but many people repeatedly interact. This research was developed by Robert Bowles and Herbert Gintis, who have suggested an analytic framework highlighting the importance of the so-called strong reciprocity for the preservation of cooperative strategies among large groups. Put schematically, strong reciprocity corresponds to the notion of paying-in-kind to anyone who behaves either in a cooperative or in a free-rider, *shirker* fashion. In a very large group of experimental studies, and with reference to very different cultural backgrounds of the participants, a tendency was clearly identifiable by most players to continue cooperating with others who also cooperate, but also—and very importantly—to punish trespassers, even involving a cost for the players who exercise punishment. Notwithstanding this individual cost, and so also the formal apparent irrationality of that conduct, consistent verification is obtainable that the diffuse presence of *strong reciprocators* in any given population works excellently as inhibitors of shirking practices. Given the fact that these are noxious not only for themselves, but mostly out of the bad examples they establish and the *cascade effects*, or *threshold-dependent* noncooperative behavior

they may induce, the necessary conclusion is that *strong reciprocity* is crucial to keep interactions in the cooperative track; otherwise, the predictable sequence of small-scale noncooperative events would imply the crumbling of the whole fabric of social cooperation.

Also very importantly, the fact that this is a cross-cultural phenomenon induces Bowles and Gintis to wonder about the existence of either a biological set of dispositions making most humans inclined toward this practice or alternatively a system of cultural values so generalized as to encompass practically all societies and cultures. This aspect refers also to a tendency in large groups, and again in cross-cultural experiments, to adopt conducts of *fair division* of payoffs in the so-called ultimatum games. If there is one player establishing the ratio of partition for a certain global payoff, and the other is limited to either accept the partition or reject it, thereby implying zero-gain for both intervenient parties, most people tend to spontaneously propose partitions somewhere between a 50–50 basis and, at the most, a 70–30 scheme. Moreover, when some players suggest a more imbalanced or *unfair* model (e.g., 90–10), the person who is presented with this offer does often tend to refuse it, mostly out of moral outrage, and indeed with some pleasure associated with punishing the misbehavior. One possible conclusion is that payoffs are relevant not in absolute terms but only comparatively, and therefore implying the appeal to notions of distributive justice.

João Carlos Graça and Rita Gomes Correia

See also Complex Systems; Prediction; Rationality; Speculation

Further Readings

- Axelrod, R. M. (1997). *The complexity of cooperation: Agent-based models of competition and collaboration*. Princeton University Press.
- Axelrod, R. M. (2006). *The evolution of cooperation* (Revised edition, with foreword by Richard Dawkins). (Original work published 1984)
- Moore, O., & Anderson, A. R. (1962). The structure of personality. *Review of Metaphysics*, 16(2), 212–236.
- Vilela, M. R. (2004). Network structure of strong reciprocity. *Advances in Complex Systems*, 7(3 & 4),

357–368. <http://label2.ist.utl.pt/vilela/Papers/recipronet.pdf>

- Von, N. J., & Oskar, M. (2007). *Theory of games and economic behavior* (60th Anniversary edition, introduction by Harold Kuhn, afterword by Ariel Rubinstein). Princeton University Press. (Original work published 1955)

GENERALIZATION

The term *generalization* is often used in the literature on research and methods when we make a theoretical claim and observe the data. But do we really observe all the data and then deduce our conclusion? Roland Barthes states that the world's narratives are numerous and further explains that all narratives share structural features. Is he suggesting a generalization that applies to narratives worldwide? Or when we reach toward meaningfulness without studying all contexts; is it generalization that supports our observations?

Generalization is a necessary component in survey and research for the formation of theories, both scientific and otherwise. But how is something generalized? For example, if the first pen I pull from my bag is green, and the second pen I take from my bag is also green, can I then generalize and say that *all* the pens in my bag are green? Is generalization more in line with inductive reasoning? It appears that the validity of one's conclusion is based on some premises, contexts, and the observer's prior experiences. This does not mean, however, that generalizing is a subjective activity.

Generalizing statements constitute the explanation schema that is the method of scientific reasoning. It is used to label many aspects of research process.

The ability to express a generalization in words presupposes a number of capacities, previously realized, for intelligently using words, phrases, and clauses that make up the sentence, but generalization is observational rather than judgmental. The generalizing claims are explicit in survey or experimental research, whereas they are less explicit in qualitative research. There are two basic groups of phenomena with which the term is ordinarily linked: The process of generalization is

a *transition* from a description of the properties of a particular object to finding and singling them out in an entire class of similar objects. Here, the researcher finds and singles out certain stable, recurring properties of these objects. For example, Chinese pronounce /r/, a trill sound, as lateral /l/, so a linguist can generalize that wherever there is a /r/ sound in a loan word, it will be pronounced as /l/ by Chinese speakers.

In marking the result of the generalizing process, the researchers' ability to abstract the findings from certain varying and particular attributes of an object is recognized. For instance, a linguistics field researcher finds out that a three-language formula has been adopted in Indian primary and secondary education. If he interviews any student studying in Grade 6, then he can generalize and deduce: "Indian students know more than one language," or "The students are multilingual." Moreover, when the researcher encounters a child having no formal education, this might make him predict "The child might not know more than one language."

During the generalizing process what takes place is as follows: First, a researcher finds a certain invariant in an array of objects and their properties, and a classification of the invariant by a description; second, the researcher finds a usage of the variant that has been cast out to allow the identification of objects in a given assortment.

Developing a generalization is regarded as one of the principal processes of constructing a concept. In texts and reference books on various subjects, the teaching material, generally, is organized, so that the learners' interaction with it will lead to the relevant generalizations. The verification of the level of generalization that the learners have reached is described in the methods manual that directs the instructors in teaching.

For instance, in the English course, there is study of the morphological structure of a word. Specifically, the learners become familiar with the root of the word, with the infinitive, prefixes, suffixes, and so on. According to their textbook, the children do the following exercise. First, they copy certain texts and they find and single out the words that have a common part in them and underline it (play, plays, played, playful, playfully, playboy, etc.). Then, they recall the instances of similar lexicon that are specifically pointed out

in the textbook. These lexicons are related in meaning; they have a common root. After this, they become familiar with the definition of stem and base: "The part of a word that is left after you remove the part that changes when forming a plural, past tense, and so on." The learners again do a series of activities on finding and grouping words that have a shared root. The whole activity is known as a generalizing process to construct a concept.

Arriving at this generalization scheme in the teaching process presupposes a variety of special conditions. Chiefly, a collection of concrete impressions or a set of particular objects is needed. They provide the raw material for making a comparison; for this, the similar, common, collectively held qualities of the objects are detected. However, the collections of raw material should contain very different variants of the combination of similar qualities and should be sufficiently diversified with concomitant attributes.

Thus, to form the generalization that determines the concept of a suffix, the sets provided for comparison should contain words having the same root with different suffixes and ones having different roots with the same suffixes, words belonging to different grammatical relations, and so forth.

The adequacy and completeness of the generalization is decided by the broadness of the variations of the attributes that are associated on the existence in the raw material of extremely unusual and unexpected combinations of the common quality with the belonging form or attributes of expression—for instance, when researchers are constructing a generalization related to the concept they see, and they construct it with highly varied correlations.

Generalization is also considered as a *rule* that is inseparably linked to the operation of abstracting. Laying out a certain quality as a shared one includes separating it from other qualities. This allows the researcher to convert the general quality into a particular and independent object of the following actions. To know what is common is the consequence of having compared and recorded the common element in an object of inquiry, which is always abstract or conceivable. The exclusive separation, the finding out what is common, and its nearness to the particular—this is

the process of abstraction. Hence, when the generalization “the number 10 is 2 more than the number 8” is being constituted, there is not only an isolation of this feature from the other properties, such as size, material, and word, but also a finding out of the similar feature that for any objects 10 is two more object than eight, as a result of which the provided relationship starts to be considered as a relationship of abstract numbers, as a specific object of attention, abstracted from concrete objects.

The Logic of Generalization

In a generalizing process, a researcher discerns a claim that what is the case in one time or place will be so elsewhere on in another time. Regular and routine social life is dependent upon the progress of subjects accomplishing this. Inductive generalization is reliable in everyday life for the purpose it allows, and it includes a great amount of incorrect inductive assumption without considering the consequences. In the social sciences, as well as in the natural sciences, the researcher aspires to more reliability, which leads to the clarification of generalization toward inductive reasoning based on axiomatic (or deductive) reasoning or probability.

The physical world, particularly at the biological level, and much of the systems of the social world are constitutionally complex and characterized by feedback mechanisms; deductive generalizations are available where invariant laws function in large-scale physical systems. In social world systems, specifically, meaningful behavior is a result of consciousness, and the same circumstances or actions can generate different subjective meanings, or, contrarily, subjects may act on the basis of similarly expressed meanings. Social critics of generalization maintain that because of this, generalization is not possible.

In qualitative research, more recognizable and probabilistic reasoning may be used under certain circumstances. This generalizing form is applied from sample to universe in survey research depending upon the form of probability sampling, through which within every case, each layer has the same known probability of selection. Other strategies, such as quota sampling and multiple captures, aim to emulate a probability sample where such

sampling is impossible. In the social world, statistical generalization that is not possible concerning the issue of variance gives way to the second class of generalization.

Independent of statistical generalization is *moderatum* (moderate) generalization. Although it is possible to express *moderatum* generalizations in formal terms, yet they look like the everyday generalizations of life in their scope and nature. They are called moderate because they are open to change and they do not make all-encompassing sociological statements. For this reason, they can never lead to axiomatic generalizations because their alteration, by its nature, makes them hypothetical. Moreover, they are testable statements that might be refuted or confirmed at later stage.

It has been observed that variables such as system, subject, and task affect performance. Statistical generalization rigorously limits the generality of the results in the research. Strict attachment to the same experimental conditions, and the same fixed levels of the independent variables, would make a research project ineffective, while a reasonable process of generalization could make a research effort more productive. Generalizations can be emulated from one type of task to a similar type of task, from one class of users to a similar class of users, and from one set of experimental conditions to a similar set of conditions. But a researcher should be cautious of generalizing results when the agent upon which research is conducted varies greatly.

The classes of users, the task, and environmental and system variables are to be considered in generalization. The user classes vary in experience (knowledge, understanding, automaticity, etc.), skills (e.g., writing, typing, clerical, motor, visual, drawing), personality (temperament, style), and motivation (excitation level, achievement). The task conditions show variation in demands (time pressure, cost of errors, etc.), definition (problem definition, types of targets, restrictions on solution, etc.), and knowledge domain (e.g., structure; concreteness, richness). The system conditions vary in input devices such as mouse and keyboard and display devices (monitor and audio). The environmental conditions show variation in social and organizational (role, status, and audience) and physical (sound, lighting, etc.).

Generalization in Quantitative Research

According to the *Cambridge Advanced Learner's Dictionary* (2011), a generalization is “a written or spoken comment in which you say or write something very basic, based on limited facts, that is partly or sometimes true, but not always, or when someone generalizes by using such statements.” Consequently, *generalization* is a contentious word. However, there is a consensus that quantitative researchers tend to generalize inferences and findings from a model statistical sample from which the data were collected. Such generalization is known as *statistical generalization*. For getting model statistical samples, chiefly random sampling techniques are used, such as systematic random sampling, stratified random sampling, cluster random sampling, and simple random sampling. However, this uncontrolled use of non-random samples has resulted in null hypothesis significance tests, and consequently, most of the data in behavioral and social sciences remain undistributed. Quantitative research also uses relatively small sample size, and the finding of the small population is generalized to the large population. This calls into question the extent to which interpretations, results, and findings can be justifiably generalized.

Generalization in Qualitative Research

The concept of generalization is also not free from controversy. Researchers believe that approximate generalizations can be made, such as naturalistic generalization, generalizations as representing replication logic, and so on, through qualitative research. They believe that the goal of qualitative research is to gain understanding of particular social, individual, familial, and educational practices and processes that occur within a specific context and location.

In qualitative research, the generalizations are chiefly of two types: case-to-case transfer and analytic generalizations. *Case-to-case generalizations* can be emulated from one type of task to a similar type of task, from one class of users to similar class of users, and from one set of experimental conditions to a similar set of conditions. *Analytic generalizations* address theory more broadly on the basis of how selected cases or layers fit with

general constructs. These are also known as *internal generalization*. Some critics, however, have claimed that qualitative researchers may be prone to error in using a small sample as a basis to generalize their findings in thematic representations. On those grounds, analytic generalization may be criticized as unscientific and therefore contested: How can a single case become the basis of generalization? Robert K. Yin, a specialist in case study design and qualitative research, explains that analytic generalization (internal generalization) happens in two steps: First, the generalization or finding bears upon a specific theory, and second, this theory may be applied in identifying situations in which similar things might happen. For example, the linguist Noam Chomsky's innateness hypothesis holds that at least some knowledge about language exists in humans at birth, and on that basis, Michael Tomasello, a scholar of language acquisition, claims that children have specific cognitive capacities at birth that promote linguistic competence. Further, the psychologist Steven Pinker argues that humans have an innate language instinct. Yin argues that internal generalization does not represent a sample; instead, the goal is to expand and generalize theories.

Amitabh Vikram Dwivedi

See also Deduction; Induction; Inference; Knowledge; Reasoning

Further Readings

- Creswell, J. W. (2003). *Research design: Qualitative, quantitative, and mixed method approaches*. SAGE Publications.
- Denzin, N. K., & Lincoln, Y. S. (2005). *The SAGE handbook of qualitative research*. SAGE Publications.
- Filstead, W. J. (1970). *Qualitative methodology: Firsthand involvement with the social world*. Markham.
- Neuman, W. L. (1997). *Social research methods: Qualitative and quantitative approaches*. Allyn and Bacon.
- Thomas, R. M. (2003). *Blending qualitative and quantitative research methods in theses and dissertations*. Corwin Press.
- Trask, R. L. (1999). *Key concepts in language and linguistics* (2 ed.). Routledge.
- Whitehead, F. (1992). Roland Barthes's narratology. *The Cambridge Quarterly*, 21(1), 41–64. Retrieved from

December 14, 2020, <http://www.jstor.org/stable/42971737>

Yin, R. (1994). *Case study research: Design and methods*. SAGE Publications.

Yin, R. (2010). Analytic generalization. In A. J. Mills, G. Durepos & E. Wiebe (Eds.), *Encyclopedia of case study research* (pp. 21–23). SAGE Publications.

GENETIC DRIFT

Genetic drift is a process whereby a trait (such as a gene, a protein's specific amino acid sequence, or a phenotypic feature) becomes more or less common in a population. Other ways by which trait frequencies change are mutation, migration, and natural selection. A question that biologists commonly try to settle is whether a certain trait is prevalent in some population because there has been selection for it (because it enhances the likelihood of survival and reproduction) or because there has been selection of it but not for it (because it does nothing to enhance the likelihood or survival and reproduction, but it is linked to traits that have been selected for) or instead because of genetic drift. The importance of drift in the history of life and its precise difference from selection remain objects of research.

Varieties of Drift

One kind of change in trait frequencies that is commonly classified as drift is “indiscriminate parent sampling.” Suppose there is a population consisting of red squirrels and brown squirrels, for which color is a heritable trait. A sudden natural disaster (such as a fire, epidemic, or volcanic eruption) occurs, killing a large fraction of the squirrels. As it happens, a larger fraction of the red squirrels die than do the brown squirrels—not because of any fire-resistance associated with the brown color, but by chance. For instance, perhaps a larger fraction of the red squirrels happened to be in the path of the fire. Consequently, a greater fraction of the brown squirrels than the red squirrels become parents of the next generation. The resulting change in the frequencies of redness and brownness in the population is a result of drift, not natural selection.

In this example, the change in trait frequencies was random in that no difference in fitness between the red and brown squirrels was responsible for it. (For this reason, genetic drift is sometimes called “random drift.”) But although random in this respect, the change in trait frequencies is not inexplicable or the result of irreducible, fundamental randomness in nature (as in the decay of a radioactive atom). For each red squirrel who died, there were causes of its death, and those causes may even have made its survival impossible.

Another kind of change in trait frequency that is commonly classified as drift is the *founder effect*. Suppose that a lake is populated by a given species of fish. Most of these fish have spiky fins, but some have smooth fins instead. Suddenly, a landslide occurs, permanently cutting off one small part of the lake. The fish who were in that part when it was isolated will be the founders of a new population. These founding fish may not constitute a representative sample of the original population. For instance, perhaps by chance all of them have spiky fins—or perhaps the majority have smooth fins. This change in trait frequency is the product of drift, not selection.

This example illustrates that drift can occur even if the various alternative traits that are changing in frequency have different levels of fitness. Perhaps having smooth fins is more effective for swimming than having spiky fins, so having smooth fins instead of spiky fins enhances a fish's fitness. This selective advantage, however, played no role in the frequency change. If having smooth fins was more common in the newly isolated population than it was in the original population, this frequency change was not the result of smooth fins enhancing fitness.

Although drift can occur among traits that differ in fitness, there are thought to be many instances of drift where the traits changing in frequency are equally fit. The individuals in a given population may all have a given protein molecule, but there may be differences among them in the specific amino acid sequence of that protein. These differences may not affect the protein's capacity to perform its biological function, the individual's vulnerability to disease, and so on. In that case, these differences are invisible to selection. Any change in the frequencies of these amino acid sequences (absent mutation and migration) results from drift.

The fish example can also illustrate that drift tends to have a greater impact on trait frequencies insofar as the population is smaller. To the extent that the founding population is a smaller sample taken from the original population, it is to that extent less likely to be a representative sample. The frequencies of spiky fins and smooth fins in the founding population are less likely to match their frequencies in the original population.

An analogy to drift is the random sampling of marbles from an urn. Suppose 30 of the 100 marbles in some urn are yellow and 70 are blue. Suppose I reach into the urn without looking and remove 20 marbles at random. (These marbles are analogous to the founding population of fish.) The proportion of the removed marbles that are yellow may fail to approximate the 30% of yellow marbles in the original population. But if I had removed only five marbles instead of 20, the probability would have been much higher that my sample would have failed to be representative of the original population.

On the other hand, suppose (unrealistically) that the urn contained infinitely many marbles and that an infinitely large sample was drawn randomly from the urn (leaving infinitely many marbles behind). Then, the likelihood of the sample matching the frequencies of yellow and blue marbles in the urn will be arbitrarily high. For this reason, drift is vanishingly weak when the population size is infinite. Of course, no real urn or living population is infinitely large. But biologists sometimes use infinite models to simulate the development of populations. In these models, drift is ignored.

Another kind of case that is commonly classified as drift occurs when a population of sexually reproducing organisms produces a new generation. Some of the parents will contain a gene for one trait and a gene for an alternative to that trait, as when a human being with brown eyes has one gene for brown eyes and one gene for blue eyes. Each gamete will contain only one of these genes. It could happen that among these heterozygote parents, the fraction of their gametes containing the brown-eye gene turns out by chance to be greater than the fraction containing the blue-eye gene. Likewise, even if the two genes are equally represented in the gametes, it could happen by chance that a larger fraction of the brown-eye gametes than of the blue-eye gametes contribute to the next generation. The resulting change in the

population's trait (and gene) frequencies would then be the product of drift, not selection.

The Distinction Between Drift and Selection

In the preceding examples, drift clearly occurs and selection clearly does not. However, there may be many realistic cases where the same trait is subject to both drift and selection or where it is unclear how the case should be characterized.

Suppose that in a population of moths, some are dark and the others are light. Suppose that 60% of the trees have dark-colored bark and 40% have light-colored bark. Dark moths blend in when they land on dark-colored trees but are easier for predators to spot when they land on light-colored trees, and the reverse holds for light moths. Suppose that this difference in camouflage is the only factor at work in selecting for moth color. Since there are more dark than light trees, the dark color in moths is fitter than the light color. However, suppose that during some period of time, the frequency of light moths increases. Although the light moths during that period landed proportionately on light and dark trees, the dark moths happened by chance to land disproportionately on the light trees and so tended to be eaten more readily by predators. Is this change in the moth population's trait frequency a result of selection or drift?

Drift appears to be operating because it was just by chance that during this period, dark moths landed disproportionately on light trees. But selection also appears to be operating, since on a light tree, dark moths are at a selective disadvantage. Moth color makes a difference to the likelihood of predation and predation was responsible for the change in color frequency. The sampling of moths was not indiscriminate; moth color, which is responsible for moth fitness, made a difference.

Although the dark moths were fitter than the light moths, the light moths increased in frequency. This does not mean that natural selection was not operating. Selection is a statistical process. A fitter trait is more likely than a less fit trait to increase in frequency, but it is not certain to increase in frequency. Unlikely outcomes (like a long run of heads from a fair coin) sometimes happen. That the outcome was unlikely does not mean that the likelihoods (in our case, the fitnesses) were not responsible for the outcome. An improbable result of selection is no less a result of selection than a

more probable result of selection (though it may be more difficult to ascertain the traits' relative fitnesses if their fitnesses make the result unlikely).

What, then, distinguishes a result explained by selection from a result explained by drift—or are some outcomes explained by both? One suggestion is that selection explains a change that the traits' fitnesses make likely whereas drift explains a change that the traits' fitnesses make unlikely. This view would construe the decline in the frequency of dark moths as drift, not selection, even though predation and moth color (which are responsible for the traits' fitnesses) produce this result. Conversely, a frequency change can have been likely, considering the trait fitnesses, and nevertheless can have been the product of drift rather than selection. For instance, suppose that a goldenrod population consists of equal numbers of plants producing smooth seeds and plants producing wrinkled seeds. Suppose that the trait of producing smooth seeds is fitter. However, a rare flood occurs, killing nearly all of the population's members. The only survivors are the few plants that were living outside of the flood zone, and by chance, more of them produce smooth seeds than wrinkled seeds. This increase in the frequency of the smooth-seed trait is just what was likely, considering the trait fitnesses. Yet this is a case of drift, not selection.

Another suggestion is that whether an outcome is explained by selection or drift is not determined by whether the outcome is caused by the factors responsible for the traits' fitnesses. Rather, an outcome is explained by drift exactly when it is a characteristic manifestation of the fact that selection is a statistical (i.e., a chancy) process. That the light moths increase in frequency, even though the dark moths are fitter, is a product of drift in that this outcome is made possible by the fact that selection is a statistical process. Population biology likewise treats the occurrences of rare population bottlenecks (such as the flood that nearly wiped out the goldenrod population or the small founding population of fish) as statistically freak events—flukes that are bound to happen from time to time in the long run but are unlikely in any short run. Similarly, that two populations starting with the same trait frequencies and experiencing the same selection pressures (and no mutations and migrations) diverge in their trait frequencies is explained by drift not because the trait fitnesses are irrelevant to the final trait frequencies but because this divergence is characteristic of

statistical processes. It is not as if one of these two populations experienced only selection and the other experienced only drift. The difference between them is explained by drift even though selection was operating on both populations.

Marc Lange

See also Biology, Evolutionary; Life Sciences, Contemporary; Modeling; Natural Selection; Philosophy of Science

Further Readings

- Beatty, J. (1984). Chance and natural selection. *Philosophy of Science*, 51(2), 183–211.
- Hodge, M. J. S. (1990). Natural selection as a causal, empirical, and probabilistic theory. In L. Kruger, G. Gigerenzer, & M. S. Morgan (Eds.), *The probabilistic revolution, 2, Ideas in the sciences* (pp. 233–270). MIT Press.
- Kimura, M. (1968). *The neutral theory of molecular evolution*. Cambridge University Press.
- Lange, M. (2013). Really statistical explanations and genetic drift. *Philosophy of Science*, 80(2), 169–188. <http://www.jstor.org/stable/10.1086/670323?origin=JSTOR-pdf>
- Millstein, R. (2002). Are random drift and natural selection conceptually distinct? *Biology and Philosophy*, 17(1), 33–53. <https://doi.org/10.1023/A:1012990800358>
- Rosenberg, A. (1994). *Instrumental biology or the disunity of science*. University of Chicago Press.
- Sober, E. (1984). *The nature of selection*. Bradford Books/MIT Press.
- Wright, S. (1931). Evolution in Mendelian populations. *Genetics*, 16(1), 97–159.

GEOLOGY

Geology, Greek for *Earth study*, is exactly how it reads: the science of the Earth. This can be defined as the solid or liquid components that comprise Earth our planet, but also the science of solid and liquid structures on other celestial bodies such as the Moon or Mars. Geology uses scientific methods to explore and understand the dynamics of the planets, their resources, and their sustainability through time. One major component of geology is the introduction of deep time by 18th- and

19th-century theorists, on the basis of physical evidence as opposed to biblical chronology. Thus, geology includes the study of events that happened thousands, millions, and even billions of years ago and processes that spanned thousands, millions, and sometimes billions of years.

Geology is a complex and variable field. Geoscientists often specialize in one discipline within geology that often overlaps strongly with another scientific field (e.g., chemistry, biology, physics). A few of the major disciplines in geology are described here. *Geochemistry* is the investigation of chemical elements in rocks and minerals to help explain mechanisms behind formation, distribution, and deformation. *Geochronology* is the science of timing, duration, and rates of Earth processes and events. *Geomorphology* is the study of present-day landforms and their processes of formation. *Geophysics* is the branch of science that studies the interior properties of the Earth, often using remote sensing techniques. *Hydrogeology* is the study of the movement of water beneath the Earth's surface. *Paleontology* is the study of life on Earth. *Petrology* is the description of rocks and rock-forming minerals including processes that form and later alter them. *Resource exploration* is the discovery and exploitation of important resources (e.g., water, hydrocarbons, minerals). *Structural geology* is the study of rock deformation and map making. There are branches of science that are closely related to geology including atmospheric science, anthropology, climatology, environmental science, geotechnical engineering, and oceanography; these are all very important Earth-related sciences but are outside the scope of this entry.

The importance of geology cannot be overstated. Almost everything a person uses daily is at least partly a result of work by a geologist. Some examples of everyday rewards of geology include the gasoline used to run a vehicle, the diamonds for an engagement ring, plastic containers for storage, and metals used in construction. Furthermore, geology can be linked to certain famous works of art, and geological factors have sometimes led to wars between nations.

Structure of the Earth

When the Earth was first formed, it likely had a uniform composition from core to crust. Through time the Earth was subject to compression caused

by gravitational forces and radioactive decay of elements; these processes increased the temperature and pressure within the planet, resulting in a series of layers forming that differed by density and composition. On Earth and other planets, temperature increases with depth. For example, at or near the Earth's surface rocks are close to 0°C; temperatures increase to approximately 7,000°C at the Earth's core. The rate of temperature increase, called the *geothermal gradient*, increases approximately 30°C for every kilometer of depth. The continuous flowing of heat outward from the interior of the Earth induces convection beneath the crust; this convection is the driving force behind the movement of tectonic plates (discussed in the Plate Tectonics section of this entry).

Geoscientists describe the concentric layers that comprise the Earth by either rheology or composition. *Rheology* is the study of physical properties at different depths and pressures, and composition is the chemical makeup of Earth's layers. Both are described in detail below.

Composition of the Earth

If defined by composition, the Earth has three layers from listed from outside to inside: crust, mantle, and core. The *crust* is dominantly composed of light silicate minerals and is further separated into continental crust and oceanic crust. Continental crust underlies most landmasses and is dominantly composed of the rock *granite*, which has lower density and is thick (approximately 35 km) when compared to oceanic crust. Oceanic crust makes up most of the rock found at the bottom of the ocean floor and is dominantly composed of the rock *basalt*, which has higher density and is thin (approximately 8 km thick).

Underlying the crust is the mantle. The *mantle* is composed of high-density iron- and magnesium-silicate rocks. The mantle extends to a depth of 2,885 km. Beneath the mantle is the iron-, nickel-, and sulfur-rich core. The core extends to the center of the Earth, which is 6,371 km deep.

Rheology of the Earth

According to rheology, the Earth can be separated into five layers listed from outside to inside: lithosphere, asthenosphere, mesosphere, outer core, and inner core. The *lithosphere* is cool, solid,

and brittle. It extends from the surface to an average depth of 100 km. If correlated to the compositional layers of the Earth, the lithosphere comprises all the crust and the upper part of the mantle.

Beneath the solid lithosphere is the *asthenosphere*. The asthenosphere behaves in a plastic manner, which means that it is capable of flow when gradual forces act upon it. The asthenosphere extends from the lithosphere at 100 km depth to approximately 700 km depth. If correlated to the compositional layers of the Earth, the base of the asthenosphere matches with the base of the upper mantle.

Beneath the plastic asthenosphere is the plastic, but more rigid, *mesosphere*. The mesosphere behaves differently from the asthenosphere in that it is subject to more extreme temperatures and pressures with increased depth. The mesosphere extends from the base of the asthenosphere (ca. 700 km) to a depth of approximately 2,885 km. If correlated to the compositional layers of the Earth, the base of the mesosphere matches with the base of the mantle.

The *outer core* is liquid and capable of flow. The outer core extends to a depth of approximately 5,155 km. Beneath the outer core is the solid *inner core*. The inner core is rigid and the most dense of all layers within the Earth. The inner core extends all the way to the center of the Earth (6,371 km deep).

How Do We Know Earth's Internal Structure?

The internal structure of the Earth has not been directly sampled. The layers within are interpreted mostly by remote-sensing techniques. Some techniques that inform us about the interior of the Earth include *seismic waves*, sound waves that propagate through the Earth after an earthquake. Many earthquakes analyzed through time have produced detailed, three-dimensional images of the Earth. Seismic waves change their speed and direction depending on the physical properties of the material through which they pass, which is how geoscientists identify the different rheological and compositional boundaries.

There are two main types of seismic waves, P waves and S waves. *P waves*, or longitudinal waves, travel in the direction of propagation and can transmit through all material properties (liquids and

solids). *S waves*, or transverse waves, travel perpendicular to the direction of propagation and are only able to travel through solids. Comparing P waves to S waves, geologists are able to determine the rheological properties of the Earth. For example, because the outer core is liquid, S waves dissipate at the contact between the mesosphere and the outer core.

Age of the Earth

The Earth is approximately 4.6 billion years old. This age has been calculated by dating rocks and minerals found on the Earth's surface. The oldest terrestrial mineral on Earth is a zircon from Western Australia; zircon contains the element uranium, which enables scientists to use the process of uranium-lead radiometric age dating (discussed below) to determine accurate ages for mineral formation. The zircon from Western Australia has been dated as 4.4 billion years old. This mineral has survived for so long because it is nearly indestructible; zircon can be weathered, eroded, and redeposited many times.

Because the Earth is so old, it may be difficult to comprehend the time it takes for geologic events to occur. For example, sediment from a river may accumulate over centuries to create a singular sandstone bed. Geologists have developed ways to segment Earth's history into temporal groupings based on the age of rocks, minerals, and the appearance/disappearance of fossils.

Geologic Time Scale

One of the most valuable contributions geologists have made to the scientific community is the Geologic Time Scale. The passage of time on Earth is divided into geochronological units that can be referred to as years before *present* or by given names. *Years before present* means since 1950; however, because geological time is usually measured in millions of years, this is an unnecessary distinction. For example, the Cambrian Period spanned from 541 to 485 million years ago (Ma).

The divisions in the Geologic Time Scale vary in range and include *eons*, *eras*, *periods*, *epochs*, and *stages*. Time on Earth can be separated into four eons: the Hadean from 4.6 to 4.0 billion years ago (Ga), the Archean from 4.0 to 2.5 Ga, the Proterozoic from 2.5 Ga to 541 Ma, and the Phanerozoic from 541 Ma to present. These eons

represent major changes on Earth for which geoscientists made special note. For example, the start of the Phanerozoic Eon corresponds to the beginning of recognizable life on Earth. *Eras* are large separations of time and include the common separations of time in the Paleozoic (541 to 251 Ma), Mesozoic (251 to 66 Ma), and Cenozoic (66 Ma to present). Periods, epochs, and stages of time are used mostly when examining detailed sections of rock. For example, many people recognize the Cretaceous Period as the time when dinosaurs such as *Tyrannosaurus rex* roamed the Earth.

Absolute Dating

Most rocks on Earth contain, even in trace amounts, radioactive materials. The three radioactive elements on Earth are uranium, thorium, and potassium. Radioactive elements spontaneously decay, or break apart, over the course of time into atoms of other elements. Decay commences as soon as a rock is produced and the time this decay takes is consistent, so it can be measured in a lab. A half-life is the time it takes for half of the atoms of one element to decay to other atoms. For example, uranium 235 decays to lead 207; the half-life of this decay is 704 million years. By comparing the number of each atom type in a rock, geologists can determine how long the rock has been decaying. Because decay begins immediately after a rock is formed, when decay is measured the absolute age of a rock can be determined. The process of discovering absolute ages is called *radiometric age dating* and is a powerful tool in understanding the age of the Earth and the events that occurred upon and within it.

Relative Dating

Relative dating is the practice by which rocks can be correlated and compared across vast geographic areas to determine an order of events without necessarily knowing absolute ages. This can be done at outcrop scale by following basic principles of *stratigraphy*, which is the study of layered rocks to determine the order and timing of events in Earth's history. The main principles of stratigraphy are superposition, original horizontality, lateral continuity, cross-cutting relationships, inclusions, and faunal succession. The

principle of *superposition* states that in an undeformed sequence of layered rocks, the layers at the bottom are oldest and the rocks at the top are youngest. The principle of *original horizontality* states that layers of rock are deposited horizontally with some rare exceptions (e.g., igneous dykes, slope margins). The principle of *lateral continuity* states that strata are widely distributed in all directions until they reach some barrier or travel too far from the source. The principle of *cross-cutting relationships* states that any rock or mineral that cuts across another is always younger than the host rock. The principle of *inclusions* states that fragments of another rock housed within a separate rock are always older than the host rock. Finally, the principle of *faunal succession* states that organisms undergo successive changes through time, and rocks that contain fossils can therefore be compared and characterized as younger or older. The most effective fossils are those that have a large spatial distribution but a narrow temporal range. By using these stratigraphic principles, the relative timing of events in a series of rocks can be interpreted.

Plate Tectonics

Another invaluable contribution to the science of the Earth is the theory of plate tectonics, which holds that the rigid plates made up of the rigid lithosphere move on top of the plastic asthenosphere. This theory is used to explain features on Earth such as oceans, mountain ranges, volcanoes, and earthquakes. Plate tectonic theory stemmed from observations that continents separated by oceans (e.g., South America and Africa) have complementary shapes, remains of unique animals and plants (e.g., *Cynognathus* an extinct mammal ancestor from the Triassic period) being found on all southern hemisphere landmasses, and similar sequences of rocks being described across oceans. These observations culminated in a theory that tectonic plates come together and move apart through time, thereby creating mountains, ocean basins, deep sea trenches, and hotspot volcanoes. This process of continents moving apart and coming together is called a *Wilson cycle*—named after the Canadian geophysicist J. Tuzo Wilson (1908–1993)—and takes approximately 200 million years. There are currently seven major tectonic

plates on Earth: African, Antarctic, Eurasian, Indo-Australian, North American, Pacific, and South American. These major plates are in constant motion at rates varying from a few millimeters to a few centimeters per year.

As discussed previously, the lithosphere beneath oceans is called the *oceanic crust*. Oceanic crust is dominantly composed of mafic igneous rock (e.g., basalt, gabbro) that is sourced from molten material that erupts at places where two tectonic plates move apart from each other. Oceanic crust has a density of approximately 3.0 g/cm^3 and is relatively thin compared to continental crust (approximately 8 km). Continental crust underlies the continents and is composed of dominantly less dense (2.7 g/cm^3) light-colored igneous rock (e.g., granite) that ranges in thickness from approximately 25 km to 70 km in mountainous regions. Boundaries between continental crust and oceanic crust are called *continental margins* and often extend underwater hundreds of meters to hundreds of kilometers off the coastline.

There are three main types of contacts between tectonic plates: transform, divergent, and convergent. *Transform boundaries* are found when parts of the plates move along each other without any crust being created or destroyed. Arguably, the most famous transform boundary is the San Andreas Fault in the southwestern United States. The fault associated with this boundary has moved many times during the recent geological past and is constantly monitored by the United States Geological Survey. *Divergent plate margins* are places where tectonic plates move apart from each other and new crust is created. Examples of divergent plate boundaries where two tectonic plates are currently moving apart are the mid-Atlantic ridge and the east African rift valley. *Convergent margins* are places where tectonic plates move toward each other and where crust is destroyed. These convergent margins can occur where oceanic crust meets oceanic crust, where oceanic crust meets continental crust, or where continental crust meets continental crust. These interactions result in one plate; the less dense plate overrides the more dense plate and the more dense plate goes down into the asthenosphere in a process called *subduction*. As the denser plate is forced down beneath the less dense plate, pressure and heat increase. This can melt the overlying rocks, which may then

erupt as a series of volcanoes within the overriding plate, parallel to the tectonic boundary. An example of a subduction zone where two oceanic plates approach each other and volcanoes form is that of Japan. An example of a subduction zone where an oceanic plate subducts beneath a continental plate is that of the western coast of North and South America. Finally, where two continents converge, there is often a thickening of the crust with or without associated subduction. An example of a *continent–continent collision* is that of the Himalayan mountains. The friction produced by this plate collision can cause large earthquakes that may be experienced through tsunamis thousands of kilometers away from the focus of the earthquake itself.

Rocks and Minerals

A large component of geology is the study of rocks and minerals. The Earth is made up of naturally occurring elements (e.g., hydrogen, oxygen, carbon, silicon, magnesium, iron) that combine to make minerals. A *mineral* is defined as a naturally forming, inorganic, crystalline substance with a specific chemical composition. Common minerals include quartz, feldspar, amphibole, and mica.

A *rock*, in comparison, is a combination of multiple (two or more) minerals. There are three main rock types: igneous, metamorphic, and sedimentary. The three rock types vary in how they form or are deposited. All rock types can, and do, form from the remnants of older rocks in an endless loop called the *rock cycle*.

Igneous Rocks

Igneous rocks form from molten material that has cooled and crystallized. Molten material may be lava, material that has erupted at the Earth's surface, or magma, material that has crystallized in the subsurface at depths up to a few kilometers. Rocks formed from lava are called *extrusive* igneous rocks. Extrusive rocks cool over a period of seconds to days and so have very small to invisible mineral crystals. Rocks formed in shallow magma bodies are called *intrusive* igneous rocks and rocks formed in deep magma bodies are called *plutonic* igneous rocks. Intrusive igneous rocks cool over hundreds to thousands of years and have small- to medium-sized

mineral crystals. Plutonic rocks cool over millions of years and often have very large crystals.

Molten material is classified according to the amount of silica present, which is reflected in the dominant minerals crystallized. *Felsic* magma contains 65% to 75% silica by weight; common minerals in felsic igneous rocks are light minerals like quartz and feldspar. Examples of felsic igneous rocks are granite and rhyolite. *Intermediate* magma contains 52% to 65% silica by weight; intermediate igneous rocks have a near-equal mix of light (e.g., quartz, feldspar) and dark minerals (e.g., amphibole, biotite). Examples of intermediate igneous rocks are andesite and diorite. *Mafic magma* contains 45% to 52% silica by weight; mafic igneous rocks are dark in color and commonly contain minerals like pyroxene and plagioclase feldspar. Examples of mafic igneous rocks are basalt and gabbro. *Ultramafic magma* contains less than 45% silica; ultramafic rocks are dark in color and are quite rare. Examples of ultramafic rocks include komatiite and peridotite.

Igneous rocks are formed in three different places: divergent plate boundaries, convergent plate boundaries, and hot spots. The magma produced at divergent margins often is mafic in composition. Convergent plate boundaries are places where two tectonic plates move toward each other. The magma composition at these boundaries often has higher silica content (i.e., produce felsic igneous rocks). Subduction zones are one type of convergent plate margins. A subduction zone occurs when one tectonic plate approaches another tectonic plate, and the denser plate moves underneath the less dense plate. As the denser plate descends, magma can rise to the surface and form volcanoes. Hot spots are places where molten material rises to the surface not at a plate boundary and cool very quickly. A famous example of a hot spot chain is that of the Hawaiian islands.

Sedimentary Rocks

Sedimentary rocks are very common at the Earth's surface as they are composed entirely of other rocks that have been broken down, transported, and redeposited. Sedimentary rocks are described on the basis of their physical properties including composition, color, texture, fabric, and porosity. A large part of describing sedimentary

rocks is the discussion of depositional environments. Sedimentary rocks can be formed in terrestrial environments including rivers, deltas, lakes, and deserts or in marine environments such as reefs, continental slopes, and abyssal plains. There are two broad groups of sedimentary rocks: siliciclastic and biochemical. *Siliciclastic* rocks are the residues of weathering and erosion (e.g., conglomerate, sandstone, mudrock); *biochemical* rocks are accretions of biological, biochemical, or chemical material (e.g., fossiliferous limestone, coal, evaporites).

Siliciclastic sedimentary rocks form as a result of weathering, erosion, deposition, and lithification of sediment. Weathering is the breakdown—physical, biological, or chemical—of older rock to produce smaller minerals and rock fragments. These weathered fragments are transported from one place to another in a process called *erosion*. Erosion of sediment occurs by the action of water, wind, ice, and/or gravity. Eventually sediment is laid down or deposited. If left for long periods of time, the sediment then undergoes diagenesis. Diagenesis begins with burial and lithification; *lithification* is the process of turning loose sediment into rock. Compaction and cementation, or the pressing down of overlying sediment on top of buried sediment and the gluing together of the sediment by accreting small minerals between grains, are the dominant diagenetic processes occurring soon after deposition.

Biochemical sedimentary rocks are formed by the inorganic or organic chemical precipitation of particles after chemical weathering. Unlike siliciclastic sedimentary rocks, biochemical sedimentary rocks are precipitated directly in a depositional environment without transportation. These rocks are classified chiefly on the basis of their mineral composition. For example, rock that forms by precipitation of calcite is called *limestone*, rock formed from dolomite is called *dolostone*, halite and gypsum are called *evaporites*, and inorganic silica is called *chert*. These sedimentary rocks often include *allochems*. Allochems are grains that are not formed by precipitation from seawater but are extra material, often fossils (e.g., shells, feces, cyanobacteria).

Metamorphic Rocks

Metamorphic rocks, the third main rock type, always have had a previous history as a parent

igneous, sedimentary, or even an older metamorphic rock. These parent rocks undergo a significant change resulting from high temperature and/or pressure. The change from a parent rock to a new metamorphic rock happens in one of three ways: contact with heat (i.e., contact metamorphism), burial (i.e., burial metamorphism), or plate collision (i.e., dynamothermal metamorphism). *Contact metamorphism* occurs when molten material passes near, or through, older host rocks. These host rocks are heated and transformed, recrystallizing or even producing new minerals. *Burial metamorphism* occurs when rocks are deeply buried and thus exposed to the higher temperatures and pressures that exist in the lithosphere. *Dynamothermal metamorphism* occurs when two tectonic plates converge, which produces high-pressure environments such as mountain belts.

The amount of temperature and/or pressure to which a parent rock is exposed produces new minerals, called *index minerals*, in a predictable order. Index minerals in metamorphic rocks from low grade to high grade are chlorite, muscovite, biotite, garnet, staurolite, kyanite, and sillimanite. Low-grade metamorphism, whereby rocks are exposed to relatively low temperature and pressure, occurs when the metamorphosed rock retains characteristics of the parent rock. High-grade metamorphism, whereby rocks are exposed to relatively high temperature and pressure, occurs where the newly produced rock no longer resembles the parent rock. When pressure is the more dominant process (compared with temperature), compositional bands, called *foliations*, can form as minerals reorganize (e.g., gneiss).

Common metamorphic rocks formed from layered parent rocks in order from low to high grade are slate, phyllite, schist, gneiss, and migmatite. Metamorphic rocks formed from unlayered or massive parent rocks include marble, quartzite, and hornfels.

Geological Structures

To understand the geological history of an area, geoscientists need to confidently identify rocks and minerals, but they also must be able to interpret geological structures. Structures are variations, spatial or directional, in the properties of

the Earth's crust. These structures tell us how the Earth has changed through time. For example, where one type of rock contacts another there is a geological boundary, a type of structure. Geological boundaries include bedding planes, faults, intrusive contacts, and unconformities. These different boundaries inform geoscientists about the history of an area.

Some geological structures are formed at the same time as the rocks in which they are found. These structures are called *primary structures* and include features like laminae and bedding in sedimentary rocks, or lava pillows and flow banding in igneous rocks. However, many structures are formed long after the rocks in which they are found. These structures are called *secondary structures* and include foliation in metamorphic rocks, folds, unconformities, and fractures. Secondary structures are products of deformation, which is the movement of parts of the crust relative to one another.

The Earth's crust contains primary and secondary structures nearly everywhere; a major aspect of geology is that of documenting and understanding these structures. There are three levels in the understanding of geological structures: *descriptive analysis* (i.e., the positions, orientations, sizes, shapes of structures), *kinematic analysis* (i.e., how the rock has been deformed or strained), and *dynamic analysis* (i.e., the forces, or stresses, that operated to deform the rocks to their present orientation).

A powerful representation of geological structures is a geological map. Geological maps are created by visiting rock outcrops during fieldwork and interpreting the areas between the outcrops. Geological maps are produced at a variety of scales that have variable levels of detail. For example, a 1:50,000 map (i.e., 1 cm on the map equals 500 m in the real world) is a common map scale used for regional geology, whereas a 1:1,000 map (i.e., 1 cm on the map equals 10 m in the real world) is a useful scale for resource exploration.

Life on Earth

The Earth became habitable for life approximately 4.4 billion years ago. Gases released by volcanic eruptions and extraterrestrial impacts formed a thick atmosphere around the Earth, consisting of

carbon dioxide (CO₂) and small amounts of nitrogen, water vapor, and sulfur. Once the Earth had cooled enough to have a solid crust, liquid water accumulated in basins, forming the oceans that cover much of the planet today. Evidence of life on Earth has been preserved as fossils in rocks; *paleontology* is the study of fossils as once-living organisms.

A fossil is a remnant of an organism preserved in the rock record. These remnants include parts of the original organism (body fossils), imprints or other indirect evidence of life (trace fossils), or molecular compounds produced by organisms (chemical fossils). All fossil types provide information about the organisms that made them and about the environment in which the organism lived or died.

The oldest evidence of life on Earth are stromatolites in rocks 3.5 billion years old. *Stromatolites* are domes of finely laminated sediment and microbial mats that still exist today in places such as Shark Bay, Australia. Complex lifeforms that can be identified as ancestors to groups of living organisms today arose in the Ediacaran to Cambrian Periods, approximately 635 to 530 Ma. Some examples include shelly fossils, corals, arthropods, sponges, and worms. The success of life at this time on Earth can be attributed to the following mechanisms: predation increasing ecological diversity, increase in Earth's oxygen concentration, and the increase in organic matter in oceans. From the Ediacaran onward, evidence of life on Earth increased dramatically. Most sedimentary rocks contain organic matter. Geologists use this information to reconstruct what the Earth was like through its long history.

Mo E. Snyder

See also Biology, Evolutionary; Environmental Studies; Natural Resources

Further Readings

- Benton, M. J. (Ed.). (2019). *Cowen's history of life*. Wiley.
- Frost, B. R., & Frost, C. D. (2019). *Essentials of igneous and metamorphic petrology*. Cambridge University Press.
- Fossen, H. (2016). *Structural geology*. Cambridge University Press.

Monroe, J. S., & Wicander, R. (2014). *The changing earth: Exploring geology and evolution*. Cengage Learning.

Nichols, G. (2009). *Sedimentology and stratigraphy*. Wiley.

GEOMETRY, CLASSICAL

Geometry has a special role in mathematics because of its connection with physical space. Today this might not be the only force, or the driving force, behind the development of new geometries, but in the ancient world it was an important aspect. Geometry was considered an access point to deeper insights about how the world is constructed and how it connected to our reasoning.

Geometry can be thought about in many different ways. We know this from today's geometries, which can be quite challenging for our experiences, which are derived from a normal three-dimensional perception of the universe. There were also different ways to develop geometry in, for example, ancient Greece, but one of them came to play a dominant role—namely Euclidean geometry. This geometry has played a paradigmatic role in mathematical thinking, and this entry traces its development and its sources as well as its impact on modern conceptions of nature.

The first section discusses pre-Euclidean geometry and the manifold roots of geometry. Next we present the Pythagorean ideas on geometry, and in particular the idea that the world consists of numbers (or of number units). This is followed by a discussion of Plato's idea that geometric concepts belong to an eternal world of ideas. The entry continues with a presentation of Euclid's *Elements*, with reference to his definitions, axioms, and theorems. We then explain how Euclid's conception of geometry differs from Archimedes' conception, and how Euclid's conception turned into an epistemic paradigm, stating that knowledge has to be organized in an axiomatic framework. This epistemic paradigm dominated for centuries, and also influenced how Nicolaus Copernicus, Johannes Kepler, Galileo Galilei, and Isaac Newton developed their approaches toward a modern science. The entry concludes with a summary of the ontological, epistemological, sociological, and

ethical issues related to the development of classical geometry.

Pre-Euclidean Geometry

In Greek, the word *geo* means “earth” and the word *metron* means “measurement,” so an etymological interpretation of the word *geometry* could be *land measuring*. As already observed in antiquity, such an interpretation related very closely to its utility. In Egypt, whenever the Nile River went over its banks and flooded the farmland, the boundaries between the different plots of land disappeared. They then had to be restored, and this presupposed land measuring. The same processes took place in Mesopotamia whenever the rivers Euphrates and Tigris flooded the farmland. In order to complete these restorations, geometry represented a set of useful knowledge, mastered as part of some everyday life practices.

In antiquity, geometry was also brought into action with respect to impressive constructions, the Egyptian pyramids being among the most remarkable examples. Their very existence demonstrates the presence of a profound geometric knowledge necessary for both the planning and the building of the pyramids. Just to make a single and possibly almost trivial point: The accurate construction of right angles must have been well known.

Ancient cities like Babylon were architectural wonders, and so was the architecture developed in India, China, and Japan. We also need to remember the Mayan civilization in South America, which from around 750 B.C.E. created remarkable cities and, from around 500 B.C.E., unique pyramid-like constructions. For certain, the Parthenon temple on top of the Acropolis in Athens demonstrates a unique mastery of geometry. The building of the Parthenon started around 490 B.C.E., but its construction presupposed competences and techniques developed over centuries and with resources from all over the ancient world. All these architectural achievements reveal profound masteries of geometry.

Geometry not only forms part of unique architectures, but it also constitutes part of everyday life constructions like houses, dams, and ships. Such constructions draw on geometric insight, although this could well be in an implicit format.

It is the existence of this form of geometric insight that we also need to bear in mind when considering the roots of geometric knowledge.

Geometric-based achievements are not restricted to physical constructions but can also form part of theoretical constructions. Astronomic observations have been completed at many places, and different geometric interpretations of what was taking place in the sky have been formulated. Mesopotamia was known for its detailed astronomic observations. Thales (ca. 624–548 B.C.E.), at times referred to as the first Greek philosopher, is known for having predicted a solar eclipse that took place in 585 B.C.E.. This might be true; however, it is not very likely that he was able to do so based on his own observations, but more likely that he drew on astronomic insight developed in Mesopotamia. This insight did rest on a sublime mastery of geometry. Thales came from Miletus, located at the coastline of Turkey with good connections to more Eastern regions with geometrical knowledge.

From these manifestations of pre-Euclidian geometry, it should be recognized that geometry refers to a diversity of forms of knowledge and techniques, rooted in different cultures and contexts. With this diversity in mind, let us now concentrate on a particular route in the theoretical development of geometry.

Pythagorean Thoughts

According to legend, Pythagoras (ca. 570–495 B.C.E.) was born on Samos, a Greek island on where the municipality of Pythagoreio, also called *Pythagorion*, has been named after him. About the person Pythagoras, we know very little. He operated in some closed circles that adopted a range of rituals and ideas, and the Pythagorean community took on what today could be considered a religious character.

One of the ideas advocated by the Pythagoreans was that the world is built up of numbers. This claim might sound strange to us today, but numbers can be thought of as being small units, so one can see the Pythagorean claim as a kind of preliminary version of an atomic theory. In fact, such a theory was later developed by Democritus (around 460–370 B.C.E.), who stated that everything is composed of small units, and he enumerated different types of such atomistic units.

The name of Pythagoras is associated with the Pythagorean Theorem, a theorem that was already known a thousand years earlier in an Indian context. Paradoxically it brings a devastating blow to Pythagorean metaphysics. According to the theorem, the length of the hypotenuse in the right-angled triangle with the legs having the length 1 is the square root of 2. If everything consists of numbers, the length of the sides in this right-angled triangle should be composed of an exact amount of small number units. In other words, the square root of 2 should be a rational number, and these were exactly the numbers known and operated on by the Pythagoreans. However, it can be easily proved that the square root of 2 cannot be a rational number. The Pythagoreans demonstrated this, although in a way different from how we would prove it today.

The proof rocked the Pythagorean community. There does not exist a fraction generated by a number of units (what we would call a rational number) out of which a line the length of the square root of 2 can be composed. Apparently then, the world cannot be composed of numbers as the Pythagoreans knew them.

Extraterrestrial Geometry and Axioms

Plato (ca. 425–348 B.C.E.) gave high esteem to mathematical insight which at the time was mostly geometric. He was fully aware of the Pythagorean insights, and he created a different metaphysics with respect to mathematical entities.

To Plato the world we experience is not composed of numbers; in fact, he argued, we cannot find any mathematical entities in this world. Such entities all belong to a world of ideas. We cannot access this world by means of our senses, only by our rationality. In our experiential world, we find many triangles, but they are not *real* triangles as we understand through formal definitions. The properties of physical triangles are only approximations of the properties of real triangles. The sum of the angles we might measure in a physical triangle may be more or less the same as the sum of two right angles, but rarely exactly so. The Pythagorean Theorem might apply more or less approximately to physical triangles with a right angle; however, the theorem only applies precisely to the real right-angled triangles, and such

triangles all belong to the world of ideas. That the Pythagorean Theorem does apply to such triangles can be proved through pure reasoning, without us doing any measuring at all in the physical world.

Aristotle (385–322 B.C.E.) presented a general format for how any body of knowledge should be presented, be it with respect to biology, ethics, or anything else. As we will see in a moment, Euclid's geometry became the quintessential example of this axiomatic approach. Aristotle argued that a science needs to start with definitions of the necessary concepts. These would be followed by a list of axioms, and from these axioms the body of the theory should be developed through logical deductions. Aristotle had developed what at the time was considered the only imaginable logic, and he knew exactly how many valid argument structures existed. In this way, Aristotle outlined the general structure of knowledge as represented in an axiomatic format.

Euclid's *Elements*

In the *Elements*, Euclid (ca. 325–265 B.C.E.) presented an axiomatization of geometry. Aristotle formulated his knowledge paradigm well before Euclid brought together the *Elements*; but long before Euclid did so, there had been many attempts to produce a set of axiomatizations in geometry. The one presented by Euclid did draw on these previous efforts and became the most coherent.

The *Elements* starts out with some definitions; thus, a point becomes defined as that which cannot be divided. Then follow five axioms, which should be understood as the construction techniques allowed within the axiomatic system:

1. To draw a straight line from any point to any point.
2. To produce (extend) a finite straight line continuously in a straight line.
3. To describe a circle with any center and radius.
4. That all right angles equal one another.
5. That, if a straight line falling on two straight lines makes the interior angles on the same side less than two right angles, the two straight lines, if produced indefinitely, meet on that side on which the angles are less than the two right angles.

The fifth axiom can be given different formulations; one such is:

6. Through a given point P outside a given line L, there can be drawn one and only one line parallel to L.

With these axioms as the starting point, the deduction starts, and theorem after theorem is proved in the 13 volumes that constitute Euclid's *Elements*. Geometric constructions are well-defined procedures whereby only two instruments are allowed: the compass and the ruler. Constructions should be made by circles and straight lines, and only by these two simple geometric figures. The first theorem in the *Elements* states that an equilateral triangle can be constructed.

The proof that a certain geometric figure can be constructed is similar to showing that this figure does exist in the world of ideas, at least according to Plato's interpretation. It also turned out that certain constructions are difficult, like the trisection of an angle. Later it was proved that it was in fact impossible to make this trisection by means of compass and ruler.

Euclid concludes the *Elements* by showing that it is possible to construct the so-called Platonic solids: the tetrahedron, cube, octahedron, dodecahedron, and icosahedron. These *regular* and *convex polyhedrons* had been investigated carefully long before Euclid did so, and it was well known that the five Platonic solids were the only possible regular solids. In finishing his exposition by investigating the Platonic solids and how they can be constructed, Euclid reached a powerful conclusion of a study that was initiated by showing the construction of an equilateral triangle.

We do not know much about the person that was Euclid. He worked in the Egyptian city of Alexandria and he wrote in Greek, the academic language of the time. However, he need not have been born in Greece but could have come from Egypt, from Mesopotamia, or from many other places, as the academic world at that time involved travel and interaction.

Geometric Paradigms

The *Elements* set out a paradigm for geometric thinking that turned into a paradigm for how to organize knowledge in general. This paradigm stayed dominant for more than two thousand

years, although it was not the only conception of geometry that was established.

Archimedes (ca. 287–212 B.C.E.), living much of his time in Alexandria, contributed profoundly to geometry; however, he did not operate within the Euclidian paradigm. Archimedes showed how a trisection of an angle could be constructed. In doing so Archimedes used a spiral, which according to Euclidian measures was not allowed. It was simply not part of the axioms of the Euclidean system, and it shows how values and normative decisions played a significant role in the otherwise clean and logical birth of Euclidean geometry. The straight line and the circular motion were considered the eternal movements in the Platonic approach and were hence chosen to be the foundation of proper constructions in geometry.

By operating with infinitesimals, Archimedes went far beyond the deductive pattern presented in the *Elements*. By doing so, he was for instance able to calculate the exact volume of a surface and of a sphere. Archimedes was preempting the reasoning much later applied in calculus and differential geometry, but such reasoning remains foreign to the Euclidean paradigm and was therefore not handed down from generation to generation in the training of practitioners of geometry.

Throughout antiquity, geometry was brought into action for planning and completing architectural constructions: building temples, aqueducts, bridges, dams. Vitruvius (ca. 85–20 B.C.E.) wrote the book *On Architecture* (*De Architectura*), which appeared in 10 volumes. He highlighted the importance of geometry for completing any architectural construction. In *On Architecture*, we are dealing with a conception of geometry quite different from the one set by the Euclidian paradigm. We are dealing with what later became referred to as *descriptive geometry*. One important challenge to descriptive geometry is to provide two-dimensional descriptions of three-dimensional objects. This branch of geometry has been applied during all subsequent historical periods and became further refined as part of the Industrial Revolution, when many new machines had to be described before actually being constructed. In German, descriptive geometry is called *darstellende Geometrie*, where the word *darstellende* can be interpreted as “performing” or “manufacturing.” This etymological interpretation highlights that through geometry

one *performs* a possible construction—a feature of geometry that was anticipated by Vitruvius.

Geometries of the World

Geometry has been widely used for formulating worldviews about the cosmos and the positioning of sun, earth, and planets. This connection continues the tradition from the Pythagoreans, Plato, and Euclid to connect geometry to the laws governing physical space on an astronomical scale.

An important such geometric worldview was formulated by Ptolemy (around 100–170), who lived in Alexandria. This view acknowledges the Aristotelian outlook by placing the earth at the center of the universe and by considering all heavenly movements to be circles.

Ptolemy had access to a huge number of astronomical observations, meaning that the actual movements of the planets in the sky were well documented. Unfortunately, however, these observations did not fit any simple model of the universe with the movements of the planets presented as circles with the earth at the center. Whatever size that was looked at for the radius of the circles one assumed, a gap occurred between model and observations. Owing to fundamental metaphysical assumptions at the time, one had no alternative but to take as a given that heavenly movements were circles and to attempt to repair the model by introducing a number of epicycles. An *epicycle* is a circle with its center positioned at the periphery of another circle, so by operating with a sufficient number of epicycles as well as with other circle-preserving devices, Ptolemy managed to provide a model of the universe that, somehow, was in accordance with the astronomical observations of the time.

During the whole medieval period, Ptolemy's geometric descriptions of the universe dominated thinking in the Western world. The Catholic Church declared this description to be in accordance with religious belief, such that by reference to the Bible, one could demonstrate that the earth was positioned in the center of the universe.

Observations of the movements of the planets became more and more precise, requiring that the geocentric model of the universe needed further geometrical adjustments by way of increasing the number of epicycles. After more than a thousand years of adjustments, the Ptolemaic earth-in-center

world model lost all traces of elegance and simplicity and became somewhat of a geometrical monster of complexity.

It was only when *On the Revolutions of the Heavenly Sphere* (*De Revolutionibus Orbium Coelestium*) by Nicolaus Copernicus (1473–1543) was published in 1543 that a new geometric description of the universe was presented in an elaborated format. Copernicus was a superb geometrician, and he showed that by locating the sun in the center of the universe and the earth and planets in circles around the sun, he could considerably reduce the number of epicycles necessary for making the model fit the astronomic observations.

To place the sun in the center of the universe was a heretical act in the eyes of the Christian church. *On the Revolutions of the Heavenly Sphere* was in preparation for a very long time, and it was only when Copernicus was lying on his deathbed that the book was published. It contained a preface by *Andreas Osiander*, who explained that the model was only intended for facilitating calendar calculations, and not meant as an exposition of how the universe is in fact operating. Copernicus's model was more precise than any previous geometric model applied to heavenly bodies. However, it is doubtful whether Copernicus interpreted his model in the way Osiander indicated; he very likely had found that the sun was in fact positioned in the center of the universe, and this soon became the general interpretation of the message from *On the Revolutions of the Heavenly Sphere*.

Johannes Kepler (1571–1630) was deeply inspired by Copernicus but also troubled by the missing fit between Copernicus's heliocentric world model and the more accurate astronomic observations that Kepler had access to. For more than 10 years, he studied the movement of the planet Mercury, and he reached the conclusion that it did not move around the sun in a circle but rather in an ellipse. This was yet one more heretical thought, because Aristotelian cosmology, as adhered to by Catholic dogmatism, assumed heavenly movements to be circular. Formulating alternatives to this taken-for-granted assumption was a dangerous undertaking, which could lead to the person facing the Inquisition. The Inquisition was a powerful institution in the Catholic Church that operated with ruthlessness in its efforts to eliminate heretical thinking—the use of torture being one such possibility.

Kepler not only formulated the idea that the planets move around the sun in ellipses; he also formulated three laws specifying the nature of the movements. These formulations have become an example of the first clear signs of scientific reasoning entering our worldview. But in a strong concordance with the Pythagorean tradition, Kepler also presented the idea that the movements of the planets created musical harmonies. In addition, he elaborated on the distances between the sun and the different planets, and for doing so he made use of the Platonic solids. In Kepler's cosmology, these geometric figures were nested within each other. This highlights the endurance of the Platonic–Euclidean tradition of understanding geometry; as shown earlier in this entry, these Platonic bodies were addressed by the last (and most important) theorem of the *Elements*.

Galileo Galilei (1564–1642) was dedicated to the heliocentric worldview, and as a consequence he twice had to face the Inquisition. He did not insist on his view, even though the legend tells that at the end of the process, in 1633, he added in a low voice, referring to the earth: “*And yet it moves.*” Galilei was punished, but only moderately by having to stay confined to his home.

Like many of the people who contributed to the scientific revolution, Galilei also shared many medieval ideas. He drew up a horoscope for many people, and when young he impressed the Florence Academy by demonstrating how the size of hell, as described by the poet Dante Alighieri, could be calculated. According to Galilei, hell had the volume of $1/12$ of the volume of the earth, and the shape of a cone with its top placed in the center of the earth, while the bottom circle of the cone had Jerusalem as its center. Geometry has indeed had many purposes beyond *measuring the earth*.

A Mechanical Worldview

The Scientific Revolution had been in progress for around 100 years, and still a range of questions were left unanswered. If the earth is spinning around itself, taking one turn per day, the speed of its surface must be enormous. How would a bird, leaving its nest in a tree, ever be able catch up with the nest again considering the enormous speed of the tree?

It was also a mystery to explain the very movements of the earth and the planets around the sun.

René Descartes (1696–1650) introduced the idea that there exist some enormous whirls, or vortices, which cause the rotations. In suggesting such an explanation, he tried to observe a strictly mechanical outlook. The first mechanical clockwork mechanism driven by a spring had been presented in 1430, and its mechanism became paradigmatic for how nature was operating. In order to explain the movement of something, one needed to identify the mechanical connections that brought the movement into play. Cogwheels need to interact directly; if a wheel were missing, nothing moved. It was by establishing such a mechanical worldview that Descartes proposed his vortex theory, suggesting the existence of gigantic interplanetary whirlpools, or vortices.

Isaac Newton (1642–1727) provided explanations of heavenly movements of a quite different nature. In manuscripts from 1669, 1671, and 1676, he outlined the principal ideas of the differential and integral calculus. However, when in 1687, he published *Mathematical Principles of Natural Philosophy* (*Philosophiæ Naturalis Principia Mathematica*), often referred to as the *Principia*, he did not make use of this mathematical insight. Instead, he carefully followed the Euclidian tradition. He started out by providing some definitions, and then he presented three laws of motion, which functioned as axioms for the subsequent deductions. From these axioms he deduced a range of theorems, including Kepler's laws for the movements of the planets. In the process of deduction, he did use geometric reasoning in a similar way as Euclid deduced his theorems. The *Principia* is carefully structured according to the Euclidian paradigm and in this way shows how the geometrical approach influenced many other areas of science centuries later.

With the *Principia*, Newton provided a new insight into nature. Since Aristotle it had been assumed that there exist two sets of natural laws, one set applicable to movements on the earth and another set applicable to heavenly movements. Newton made clear that the same set of laws are operating throughout the universe. According to Newton, the whole universe is operating as a unified mechanical system.

By operating with the notion of gravity, an attraction between physical bodies, Newton went beyond the clockwork mechanical worldview. The

idea of the existence of an attraction between physical bodies operating across the empty space is not consistent with any mechanical idea; rather it appears to be a spiritual idea, and hence it can be conceptualized as such.

Newton was a deeply religious person, spending as much time studying the Bible as for studying physics. Furthermore, Newton was attracted by different mystical movements, one being hermeticism, which might have prepared him for assuming the existence of an attraction operating across empty space. Whatever degree of mysticism that might be included in the formulation of the idea of gravity, it was to become an integral part of the mechanical worldview. However, with reference to the *Principia* the notion *mechanical* is not to be interpreted in a straightforward clockwork-like physical sense, but rather in a mathematical sense. The mechanisms defined through the natural laws are ensured through operations that can be specified mathematical but not physically.

The mathematical-mechanical worldview soon became generally celebrated in the scientific communities. It became assumed that Newton had identified the mechanism of nature, and that after Newton only some refinements of the theory were needed. As mentioned earlier, Newton's *Principia* methodologically follows the structure of Euclid's *Elements*, and one can see Newton's achievements as an almost universal extrapolation of the Euclidian paradigm.

Geometries of Empty Space

In *The Geometry* (*La Géométrie*) published in 1637, Descartes presented the basic ideas of analytic geometry. His presentation was not in any way systematic, and in *The Geometry* one does not find a straightforward presentation of a coordinate system, but it is certainly lurking behind the curtains. In any event, Descartes showed the analytic power of associating geometric figures to algebraic expressions.

Rooted in analytic geometry, the more modern version of a mathematical worldview entered the human imagination. One can imagine a three-dimensional coordinate system containing the whole universe, and the dynamics of objects in the universe can be represented by algebraic expressions—linear and differential equations in

all their complexity. On this basis, one can look into the future by means of explicit calculations, and one can read the past by similar calculations. The idea that the whole universe operates as a unified and deterministic machine achieved a sharper formulation by means of analytic geometry.

To begin with, it was accepted as a given that the universal space was a Euclidian space, in the sense that Euclidian geometry including all its possible theorems would be valid with respect to our physical space—maybe not valid with exactness as in the world of ideas, but with sufficient approximation. It was taken as axiomatic that the universe was a three-dimensional Euclidian space, a view that became so enduring that perceiving an alternative was almost heretical. Would it be possible to imagine the existence of a non-Euclidian space?

This question of an alternative to the Euclidean paradigm brings us back to the fifth axiom in the *Elements*. While the first four axioms appear simple and intuitively clear, the fifth axiom looks far more complex. One could consider the idea to instead transform the fifth axiom into a theorem? After Euclid, many attempts at proving the fifth axiom were carried out. One strategy was to formulate a negation of the fifth axiom, and then to make deductions from this negation and the four other axioms until a contradiction would appear. Certainly, strange consequences were deduced, and several times it was concluded that the fifth axiom was proved. However, only in the beginning of the 19th century was it fully recognized that the strange consequences were indeed not contradictions but theorems in a non-Euclidean geometry.

There now exist several different forms of non-Euclidean geometry. One can negate the fifth axiom by assuming that there does not exist any parallel line going through a point outside a given line. This leads to what is called *elliptic non-Euclidean geometry*. One can also assume that there exist more than one parallel line going through a point outside a given line. This leads to what is called *hyperbolic non-Euclidean geometry*. It was the elliptic version that Albert Einstein used when formulating the theory of relativity. In this way, he described the universal space as being non-Euclidean and by doing so established a profound paradigmatic shift in our conception of nature. With these remarks, however, we have moved far beyond classical geometry, whose roots

nonetheless still play a significant role in our perception of reality and space.

Four Dimensions of a Philosophy of Geometry

In the book *Connecting Humans to Equations: A Reinterpretation of the Philosophy of Mathematics*, Ole Ravn and Ole Skovmose suggest that a philosophy of mathematics operates with four dimensions in being ready to address ontological, epistemological, sociological, and ethical issue. Let us now explore these four dimensions with respect to geometry, and in this way reflect on the presentation of classic geometry that appears in the previous sections.

Ontology concerns what is assumed to exist, and with respect to geometry it raises questions like: What is a point? What is a line? What is a plane? And more generally: What is the nature of geometric entities? The Pythagoreans claimed that numbers had a real and independent existence from human beings, and that the world was composed of number units. This audacious ontological claim was challenged by the discovery of irrationalities. Plato redirected the ontological discussion by inventing a world of ideas that hosted geometric and other mathematical entities in a realm of their own.

A more modern and nihilistic ontological account of geometry is found in Hilbert's *Foundations of Geometry*, first published in German in 1899. At that time, several mathematicians had expressed doubt about the validity of the deductions presented in the *Elements*. It was pointed out that Euclid, when proving the theorems, had used a range of implicit assumptions besides the explicit formulated five axioms. In the *Foundations of Geometry*, Hilbert made a reformulation of the start of Euclidian geometry by presenting not five but 20 axioms. This completed system represents a kind of perfection of the *Elements* and it goes to show again one final time that the impact of Euclid's geometry on all later geometries cannot be overstated.

As a consequence of this perfection, Hilbert could announce that for proving geometric theorems the nature of notions like point, line, and plane are irrelevant. The only thing that counts in geometry is the interrelationships between these

entities as stated in the axioms. Instead of points, lines, and planes, one could as well talk about beer mugs, tables, and chairs. By making this claim, Hilbert presents the ontological issue as being irrelevant to geometry. As part of the broader formalist program in mathematics, as Hilbert later advocated, ontological issues are irrelevant not only to geometry but to mathematics in general.

Epistemological issues concern how we obtain geometric knowledge. One common sense idea would be that such knowledge is based on empirical observations. However, the Euclidean paradigm sets a specific epistemic standard: geometric knowledge is built upon self-evident axioms, and from these axioms one deduces consequences. The logical nature of the deductions ensures that if the premises are true, then the consequences will be true as well. Because the axioms are chosen in such a way that their truths can be grasped immediately, truths will flow from the axioms throughout the whole deductive system and ensure the truth of each and every theorem.

This epistemic paradigm came to dominate, and it was assumed that not only geometric knowledge but knowledge in general should be developed according to this deductive pattern. In the *Principia*, Newton followed this pattern; however, it was applied far beyond mathematics and the natural sciences. In his *Ethics*, Baruch Spinoza (1632–1677) presents a theory of ethics, including a range of statements concerning how to behave. The *Ethics* starts with definitions, followed by one axioms, and from these Spinoza deduces theorem after theorem, which constitutes his theory of ethics.

By the *sociological* dimension of a philosophy of geometry, we have in mind reflections about how social factors have an impact on how geometry is formulated and perceived. One such factor is Eurocentrism. Eurocentrism in its general format has been explored by Martin Bernal in *Black Athena: Afroasiatic Roots of Classical Civilization*, and with particular reference to mathematics by George Gheverghese Joseph in *The Crest of the Peacock: The Non-European Roots of Mathematics*.

The formulation of the modern worldview took place during the period where the process of brutal colonizations was in full development, and the slave trade turned into a widespread business. These processes integrated different versions of racism as part of the common outlook in the West.

Slavery had existed throughout all historical periods, but as part of the modern outlook the idea developed that some groups of people were naturally born to be enslaved. Compared with White people, Black people and people from Asia, the Pacific, and the Americas were held to be inferior.

This outlook also had implications for how the origin of geometric thinking was interpreted. In keeping with the modern, Eurocentric perspective, it was difficult to acknowledge that geometric knowledge could emerge in the East or in Africa. And because such knowledge evidently had been developed in these places even earlier than in the European sphere, so it was argued that those were only preliminary versions of proper geometry. Such versions, it was claimed, were based on empirical insights rather than on deductive patterns, and that only when Euclid brought together the *Elements* did geometry find its proper and sublime form.

To support the claim that geometry was a Western achievement, Greek culture was interpreted as belonging to the West, and its connections to the East and to Africa were downplayed. The Eurocentric interpretation of the history of geometry and of mathematics in general became an integral part of the modern outlook and has remained dominant right up to the present time.

By the *ethical* dimension of a philosophy of geometry, we refer to reflections on the social impact that might be exercised through geometry. Let us start considering an ideological and meta-physical impact. Cosmologies have been formulated through geometry, and paradigmatic elements of the Euclidean formulation of geometry have been defining parts of such worldviews.

Euclid considered only circles and straight lines as being the basic geometric entities, and certainly not any spirals. As we saw, such paradigmatic assumptions had a huge influence on the way Ptolemy elaborated on his geocentric worldview. What other kind of movements could be considered? It was not until the time of Kepler before it was possible to question this deep impact of the Euclidean paradigm when thinking about nature.

However, geometry is enlisted in developing not only metaphysical but also physical constructions. If we pay less attention to Euclid's presentation in the *Elements* but look at geometry as applied in architecture and in engineering by means of descriptive geometry, we see powerful achievements completely

based on geometric insight. Geometry is an integral part of three-dimensional simulations of possible new constructions, whether it be it a tunnel, an airport terminal, or a car model. Three-dimensional simulations are also used in medicine, where, for example, robots can assist in performing surgery. Similar simulations also constitute an integral part of modern war technology as, for instance, in the development of guided missiles that can cross vast distances in search of their targets.

In conclusion, this entry has shown that geometry can be brought into action within many different contexts. Any such actions are in need of critical thinking about the reasoning behind them and the values imbued in these actions. Geometry in action calls for ethical reflection.

Ole Ravn and Ole Skovmose

See also Epistemology; Geometry, Non-Euclidean; Laws of Nature; Worldview

Further Readings

- Bernal, M. (1987). *Black athena: Afroasiatic roots of classical civilization, 1: The fabrication of Ancient Greece, 1785–1985*. Rutgers University Press.
- Davis, P. J., & Hersh, R. (1981). *The mathematical experience*. Birkhäuser/Houghton Mifflin.
- Davis, P. J., & Hersh, R. (1988). *Descartes' dream: The world according to mathematics*. Penguin.
- Joseph, G. G. (2000). *The crest of the peacock: The non-European roots of mathematics*. Princeton University Press.
- Joyce, D. E. (1998). *Euclid's elements*. <http://aleph0.clarku.edu/~djoyce/java/elements/elements.html>
- Ravn, O., & Skovmose, O. (2019). *Connecting humans to equations: A reinterpretation of the philosophy of mathematics*. Springer.
- Tarnas, R. (1991). *The passion of the Western mind: Understanding the ideas that have shaped our world view*. Ballantine Books.

GEOMETRY, NON-EUCLIDEAN

Other than the geometry described in the *Elements* of the ancient Greek mathematician Euclid (born c. 300 B.C.E.), the first geometry that could be capable of describing physical space was

proposed in the 1830s independently by Nikolai Lobachevskii in Russia and János Bolyai in Hungary. When it was eventually accepted in the 1850s and 1860s as a valid piece of mathematics, after its having being reworked by Carl Friedrich Gauss, Bernhard Riemann, and Eugenio Beltrami, it provoked a widespread revision of ideas about what constitutes a geometry, and by 1900 a revision of, and return to, axiomatic methods in mathematics. Its consequences also forced a rethinking of the relation of geometry to physics, and in that sense it is a progenitor of the general theory of relativity, later proposed by Albert Einstein.

Background and Origins

By the year 1800 Euclidean geometry, as described in any of several different editions of Euclid's *Elements*, was taken to be the bedrock of mathematics, along with elementary arithmetic and some simple algebra. For many educators, it was also taken to be an exemplary display of human reasoning that established, through rigorous thought alone, a true description of the universe around us, the space for Isaac Newton's celestial mechanics that he had called *God's sensorium*. Eminent philosophers such as John Locke, David Hume, and Immanuel Kant treated it as such, and even though René Descartes had proposed a different approach to geometry, he did not differ in his conclusions.

Over the course of the two millennia since it was written, commentators and editors in Greek, Islamic, medieval, and modern European traditions had found small faults with Euclid's *Elements* that they enjoyed patching up. They agreed that it was austere—it did not make clear how you might come to similar results in geometry yourself—and indeed that there were one or two substantial difficulties in the text. One, with which we shall not be concerned here, concerned the nature of irrational numbers. The other was the much-discussed matter of the parallel postulate, and that proved to be the fatal flaw in the *Elements* and the way in to other geometries entirely.

In the *Elements*, Euclid showed that in a triangle any two angles have a sum less than π . What then, of the converse case: If two lines cross a third and the angles where they cross sum to less than π , do the two lines meet, and the three lines

form a triangle? Euclid made the explicit assumption that they do, and this is the parallel postulate, and the fifth and final of his postulates in all modern editions. It has the immediate consequence that if the two lines cross a third and the angles they make sum to π , then the two lines are parallel, which is why the postulate acquired its name.

The parallel postulate is immensely useful. Once it is introduced in the *Elements*, Euclid shows that the angles sum of any triangle is π , proved that parallel lines are equidistant, and established the Pythagorean theorem. The postulate may be obscurely expressed, but the idea that there are parallel lines is highly intuitive although, unlike the other postulates, it is less than empirically certain. It is easy to show that if the angles sum to something very little less than π then the two lines will meet very far away, well beyond our galaxy. Accordingly, a tradition grew up, starting in Greek times, of attempting to prove the parallel postulate on the basis of all the other assumptions Euclid had made.

By 1800, there was a growing literature of this kind, all of it unsuccessful. The best results, many known to mathematicians of the Islamic world, only showed that, given the rest of the *Elements*, the parallel postulate was equivalent to the assumption that the sum of the angles in a triangle is π , to the assumption that parallel lines are equidistant, and to the assumption that there are similar figures, or scale copies, that are not congruent (triangles whose angles agree in pairs but whose corresponding sides are of different sizes). Euclid's assumption looked more and more sensible on pragmatic grounds—imagine a geometry in which these results are all false—but it irked some of the best mathematicians that they could not prove such an apparently obvious truth.

Three exceptions stand out. The first is the Jesuit priest Gerolamo Saccheri, who published a book in Latin in 1733, the year he died, entitled *Euclid Freed From Every Flaw*. In it tried to show that the sum of the angles in a triangle must be π by showing that any alternative was self-contradictory. He speedily showed that assuming there are triangles with an angle sum greater than π contradicted other results in the *Elements*. Then he set to work on the assumption that there might be triangles with an angle sum less than π , and he found many interesting consequences. Finally, he

claimed to have found the contradiction he sought, in what happens when two lines cross at infinity, but rightly no-one was persuaded.

The second exception is that of the Swiss mathematician and polymath Johann Heinrich Lambert (1728–1777), who wrote up an account very similar to Saccheri's but stopping short of the final claim. Instead, he said that the nature of a straight line needed to be better defined. He would be willing to agree as a matter of fact that they behaved as Euclid said they did, but he did not see why they had to do so as a matter of logic. Lambert was thus the first person to hold out the possibility that there might be another geometry that agreed in every respect with Euclidean geometry except with regard to straight lines. However, he did not publish these investigations—they were published only after his death—and he left the many theorems he had established on the assumption that the angle sum of a triangle is less than π as tantalizing indications of what a new geometry might be like. For example, the area of a triangle is proportional to π minus its angle sum.

The third exception is Gauss. In 1801, the year he turned 24, Gauss was about to establish himself as the leading number theorist and an important astronomer, and his reputation only grew after that. He entirely reformulated several branches of mathematics and opened up others, and by an accident of when he was born there was no one else like him for perhaps 20 years. The paradox in the story of non-Euclidean geometry is that Gauss published nothing on it and contented himself with a few hints in book reviews. Only after his death in 1855 was it discovered how much he had known, and even that is problematic. Still, what he did write down showed that he was doubtful that the truth of Euclidean geometry could be proved.

The Discovery of a New Geometry

Honor for the discovery of non-Euclidean geometry is famously divided between the Russian mathematician Nicolai Lobachevskii (1792–1856) and the Hungarian mathematician János Bolyai (1802–1860). Bolyai's father, Farkas Bolyai, was also a mathematician; indeed he had studied mathematics at Göttingen with Gauss, and he had picked up the bug of trying to vindicate the *Elements*, but without success. When his

son turned out to be something of a prodigy, Farkas tried to discourage him from following him, saying he had wasted his life in this work, but János did not listen. When he found that trying to defend Euclid ended in failure, however, he decided to try the opposite and explore a new geometry instead. In 1831, he published his insights in a 24-page Appendix to his father's two-volume book on geometry.

Lobachevskii's route into the subject was much the same. Finding nothing wrong with forming a different understanding of straight lines, he decided around 1829 to accept the implications of proceeding on the assumption that his new theory could also be coherent and even true. He published long articles in Russian, an article in French in the main German journal of mathematics, a short booklet in German, and toward the end of his life a book in French on the new geometry and its implications.

Both Lobachevskii and János Bolyai published strikingly similar results, so it is convenient to describe the work of just one, and Lobachevskii's is easier, so let us briefly describe it here. In plane geometry, the parallel postulate is equivalent to the statement that given a line ℓ and a point P not on ℓ there is exactly one line through P that does not meet ℓ . Lobachevskii investigated the idea that instead there might be two lines through P , one in each direction, that approach the line ℓ arbitrarily closely but do not meet it (this means that there are infinitely many lines through P that never meet ℓ , and are sometimes said to be ultra-parallel to it). The two lines that approach ℓ arbitrarily closely Lobachevskii called the parallels to ℓ through P , and they make an angle with the perpendicular from P to ℓ that depends on the length of that perpendicular: the further P is away from ℓ , the smaller the angle. Lobachevskii called this angle the angle of parallelism and regarded the length of the perpendicular as a function of the angle. By an ingenious argument he found an explicit expression for it, and by more work he found formulae connecting the sides and angles of triangles in the new geometry as a result.

These formulae bore a striking resemblance to the formulae of spherical trigonometry, which describe the geometry of a spherical Earth, and were well known to every land surveyor and map maker. Where the spherical trigonometrical formulae have sines and cosines, Lobachevskii's have hyperbolic sines and cosines (and the occasional

change of sign, as these functions require). Lobachevskii could deduce many things from these formulae. For example, very small triangles will have angle sums that are only just less than π , and so are virtually indistinguishable from Euclidean ones. This partly explains why the new geometry had gone undetected hitherto, but it also says that any test of the truth of the new geometry would have to involve astronomical measurements (very large triangles must have angle sums noticeably less than π). He suggested one such test but was unable to carry it out. The very existence of the formulae strongly suggests that the new geometry is logically sound, but Lobachevskii's derivation of them did something else. It involved working in three dimensions, and so what he proposed really could be a geometry of space.

The reception of the work of Lobachevskii and Bolyai was appalling. It was met with silence or ridicule. In Lobachevskii's case, Gauss moved to get him accepted by the Learned Society of Göttingen, but otherwise did nothing to promote his ideas other than by writing a few letters to a colleague. János Bolyai fared worse. His father sent Gauss a copy of his book, and Gauss replied to say that he knew all that János had done already, and even demonstrated a better proof of one or two of the younger man's results. János was so angry he refused ever to publish again, which naturally did not help him win the credit that was his due. Eventually, he learned of Lobachevskii's work, and was pleased both to see it and to see what he had done that Lobachevskii had not, but that was all.

The case of Gauss's reaction is even stranger. By 1831, he had indeed come to such an understanding of new two-dimensional geometry that he could say he knew all about it, and to have no doubt in his mind that it made coherent mathematical sense. Therefore, he could say that he no longer thought that Euclid's was the only possible plane geometry. But he had not investigated the three-dimensional case, as Bolyai and Lobachevskii had done, and could only say that he doubted the uniqueness of Euclid's geometry in that case, not that he had shown it to be false.

What is even stranger, is that, unlike Bolyai and Lobachevskii, Gauss did possess compelling grounds for asserting the new geometry. They had merely explored the consequences of a new definition of parallelism, they had no proof that a

contradiction did not lurk in what they had begun to explore. This greater degree of confidence came from his work on the theory of surfaces.

In the 1820s, Gauss had taken on the task of surveying the lands of Hanover and southern Denmark, and, physically and mentally demanding as that work was, he found himself led to recast the theory of surfaces. Eighteenth-century mathematicians knew some surfaces, such as the sphere and the generalizations of the ellipse, parabola, and hyperbola, but they tended to think of a surface as the surface of something, like the peel of an apple. Gauss took the view that a surface was something described by two coordinates, such as longitude and latitude on the sphere. If we take x , y , and z , axes in space, then every point on the sphere has its x , y , and z coordinates with respect to those axes, and each of x , y , and z is a function of the two latitude and longitude coordinates. Gauss now decided that a surface was anything whose x , y , and z coordinates each depend on two quantities, which we may call u and v .

Perhaps the most obvious fact about most surfaces is that they are curved. Some are curved like bowls, some are curved like saddles, some surfaces are bowl-shaped in one region and saddle-shaped in another. Gauss discovered a result so exceptional he called it an *exceptional theorem*, and as so often he was right: Its full meaning was to reshape geometry. It involves little more than the Pythagorean theorem to discover that the distance between nearby points on a surface is given by a simple formula involving three functions of the coordinates u and v , which Gauss called E , F , and G . This formula has become known as a metric. What Gauss found was that exactly how curved a surface is at a point is given by another function of the three functions E , F , and G . A complicated one, to be sure, but a formula that involves those three functions of the coordinates, and does not involve x , y , or z . This means it is intrinsic, it depends on the coordinates alone. Put anthropomorphically, two-dimensional creatures living in the surface could discover how curved the surface was without ever leaving the surface.

What has this to do with non-Euclidean geometry? At some stage in the late 1830s, Gauss discovered that the trigonometrical formulae of Bolyai and Lobachevskii could be derived from his theory of curvature, specifically for triangles

on a saddle-shaped surface of constant negative curvature. However, he did not publish this result, nor did he generalize it to three dimensions. What he knew about the new geometry was only found in his desk drawers after his death in 1855, and such was his status it changed the perception of the new geometry utterly.

It's not clear why Gauss did not publish it, and he left splendid discoveries on other subjects unpublished too, but one reason may well have been that he felt uncomfortable about it. Not about the correctness of his work, so much as the controversy it would cause. Overall, the main reason for the rejection of the work of Lobachevskii is that it ran into an orthodoxy upon which everyone from the best philosophers to the humblest schoolteacher had accepted all their lives: the validity of Euclidean geometry. Yes, the new ideas were hard to understand; yes, they were not logically compelling—but they were not contested on those grounds. They were not contested at all. A leading Russian mathematician wrote that Lobachevskii was barely fit to be a schoolteacher, and most mathematicians simply ignored his work. Only the name of Gauss made people rethink.

The Acceptance of Non-Euclidean Geometry

As it happens, the year before he died Gauss examined the post-doctoral thesis of the mathematician who did most to pick up these ideas and make compelling sense of them: Bernhard Riemann (1826–1866). As part of the work he undertook for his postdoctoral qualification, Riemann took Gauss's ideas about the geometry of surfaces and extended them to any number of dimensions. This included showing that in every dimension greater than one, a metric and an intrinsic curvature at a point can be defined. He also emphasized that any set of points with a metric is a geometry and does not need to be a subset of any other space: One simply works with the coordinates and quantities that depend on them.

What then of Euclidean geometry? It becomes just one geometry among many, distinguished by being, in each dimension, the simplest geometry with zero curvature. In two dimensions, there are many geometries with zero curvature, of which the plane and the cylinder are two examples, and

spaces whose curvatures are the same at corresponding points are locally the same; which is why old-fashioned printing worked.

By the same token, non-Euclidean geometry in each dimension is a geometry of constant negative curvature, and Riemann wrote down an appropriate metric in each case that takes a disc of that dimension and radius 1 and exhibits it as a space of constant negative curvature. But does this mean it exists? To Riemann; to Eugenio Beltrami (1835–1900), the Italian mathematician who closely followed him; to Felix Klein (1849–1925), the leading German geometer after Riemann's death; and very soon to the French mathematician Henri Poincaré (1854–1912), who made innovative use of the geometry in another branch of mathematics entirely, the answer was a simple Yes. It is an entirely self-consistent description of a geometry, and nothing more need be asked of it.

What Riemann's work did not do was show that a surface with two-dimensional non-Euclidean geometry exists as a surface in ordinary Euclidean three-dimensional geometry with the same sense of distance between its points in each case. This is unlike the familiar sphere, which is the canonical example of a geometry of constant positive curvature. Riemann, following Gauss's lead, separated the question of the existence of a geometry into two parts: the first concerns the existence of a self-consistent space (nothing more than a set of points) with a definition of the distance between those points; the second concerns the question of whether that space can be embedded in another space (say a Euclidean space) without altering the distances between points. A space of constant positive curvature can be embedded in a Euclidean space of one dimension higher; a space of constant negative curvature cannot be embedded in a Euclidean space of one dimension higher—but that is no reason to say it does not exist as a mathematical object.

The simplest description of non-Euclidean two-dimensional geometry is known as the *Poincaré disc* (Figure 1). It is a map or picture of the geometry onto the points inside a circle of radius 1. The geodesics appear as arcs of circles perpendicular to the circle that bounds the disc, and angles are represented correctly, while a simple formula describes how sizes of figures appear to shrink as they are moved away from the center.

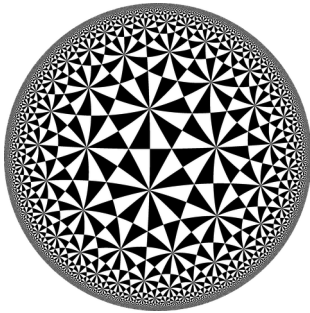


Figure 1 The Poincaré Disc Model of Non-Euclidean Geometry. All These Triangles Are Congruent, and All Their Side Arcs Are Geodesics

Source: Wikimedia Commons. Image is in the public domain.
https://commons.wikimedia.org/wiki/User:Tamfang#/media/File:H2checkers_237.png

All the seemingly paradoxical results of Saccheri and Lambert, and the definitions and theorems of Bolyai and Lobachevskii, are well displayed in the disc with a non-Euclidean metric.

What to make of all the failed attempts to exhibit a non-Euclidean geometry before Riemann? They were trying to find the appropriate surface as a subset of three-dimensional Euclidean space, and another German mathematician, David Hilbert (1862–1943), was to show in 1902 that under very modest restrictions this cannot be done.

Of course, the geometry on the sphere is also not Euclid's, and it worth observing why all the efforts to prove that Euclid's geometry was the only true geometry did not simply halt with the example of the sphere. The reason is that it departs from Euclidean geometry in two ways, not just the parallel postulate but because it is a finite space. It fails the parallel postulate because geodesics (the curves of shortest length between their points) on the sphere are great circles, and any two meet, so there are no parallels. It is finite in the sense that every geodesic on the sphere is of finite length (the circumference of the sphere) and Euclid had assumed that lines are indefinitely extendible. But spherical geometry is, nonetheless, a geometry different from Euclid's, and the reason it was never proposed in the context of a search for a non-Euclidean geometry is that, even if only tacitly, the search was for a non-Euclidean geometry that could be a geometry of physical space. For as long as it was taken for granted that space was infinite, the sphere was not a candidate. Happily, the disc

with Riemann's metric on it does have infinitely long geodesics.

Philosophers were reluctant to accept non-Euclidean geometry, especially if they were in some measure followers of Immanuel Kant. Kant had used the proof that the angle sum of a triangle is π (which uses parallel lines) as an indication of how mathematicians make constructions in order to go beyond the bare definition of mathematical objects. Various attempts were made to claim that the new ideas only established some logical result but could not make claims to be knowledge of the world, but these all lacked conviction. Gottlob Frege (1848–1925), the brilliant German logician who had trained as a mathematician, simply refused to accept non-Euclidean geometry on the singularly weak grounds that there was only one physical space and we know how to talk truly about it.

Implications for the Truth of Mathematics

Around 1900, the educated general public caught up with the new geometry and naturally wondered whether it or Euclid's geometry was true. Poincaré gave the most intriguing answer, which was that we could never know. Not because any measurements might be too delicate, but because it would be a matter of convention. Suppose you measured the angles of a large triangle whose sides are rays of light and found the sum to be less than π . There are, he said, two interpretations. One is that space is non-Euclidean and light rays are straight or geodesic, the other is that space is Euclidean and the path of a ray of light is not straight, and there is no logical way of deciding between the two. It is decided by convention among the interested parties, and simplicity will be a guide. The Poincaré disc description can be interpreted as saying that the disc displays non-Euclidean geometry and figures only appear to shrink, or that the disc displays Euclidean geometry and the figures are cooling down and shrinking. Again, Poincaré refused to say that one description was correct, but the Italian geometer Federigo Enriques (1871–1946) was among those who said that because we could not change the temperature of the disc at any point there was only one interpretation: the disc displayed a hypothetical property of space. Poincaré, however, extended his conventionalist view to any set of

circumstances where purely mathematical terms are applied to physical objects and systems and gave reasons for why we would always choose the ones that gave us the simplest theories.

The last decades of the 19th century saw an abundance of work on Riemann's rather cryptic presentation of his ideas. What gave them salience in the 20th century was Einstein's realization that the force of gravity could be understood in terms of its geodesics. This meant that it would be possible to interpret gravity as a curvature of space (more precisely, of space-time). Almost as soon as Einstein published these ideas, with the help of his mathematically better-educated friend Marcel Grossmann, another German mathematician, Karl Schwarzschild, showed that it was possible to exhibit a mathematical space that satisfied the equations Einstein and Grossmann had obtained, and the subject of cosmology was born. With the revival of cosmology after the Second World War, attention has focused on extremely massive objects which curve the space-time around them, and give it some form of positive curvature, but it is unlikely this work would have been done without the discovery of the first physically plausible non-Euclidean geometry.

The example of hyperbolic geometry did more than this. It showed that there had been something wrong all along with those claims that Euclidean geometry was true, and that a few self-evident axioms described the universe. It broke that endeavor into two: find a self-consistent mathematical system and investigate a problem about the world with an eye to seeing if some geometry is present. For a time, this even discredited the idea of starting with axioms at all. Riemann spoke only of hypotheses at the foundations of geometry, and disparaged axioms, but a generation later Hilbert showed how to give a rigorous axiomatization of plane geometry, and then of non-Euclidean geometry, and indeed of several other plane geometries that did not fit Riemann's metrical approach at all. In so doing, he established the idea that one can write down a system of axioms, check it for self-consistency, and then study any system of objects that obey the axioms. The axioms do not tell you what the objects are—they tell you how you may use them. In the same way, the rules of chess tell you how the pieces may move, but not what inner meaning they might have.

Many major branches of mathematics have since been axiomatized: much of modern algebra, geometry, topology, integration theory, probability, and others. The view became that existence in mathematics was synonymous with freedom from contradiction (a view that Frege contested in an exchange with Hilbert) and a great number of examples of resolutely pure branches of mathematics sprang up.

It would be easy to dismiss the whole enterprise as self-indulgent and unworldly, but that would be foolish. Riemann's work offered physicists a variety of universes to contemplate, and the genius of Einstein was to seize that opportunity to create his general theory of relativity. The abstract side of all of particle physics today ultimately rests on mathematically defined systems, although in the rush to understand deep physics the mathematics may well emerge in a rough and ready form.

In any case, an axiomatically defined system is one that allows you to be sure that you have not smuggled ideas into your theory that may not belong. As we enter the era of machine-checkable proofs, we can have ever-greater confidence in the validity of our arguments, which is an intriguing comparison with the much more thoroughly obscure working of machine learning and large Artificial Intelligence systems.

Hilbert defined his new geometries very often in terms of coordinates involving novel number systems, and so he began to ask about how the consistency of even the usual system of numbers can be ensured. He hoped to show that some combination of set theory and logic could do this, but in 1931 Kurt Gödel showed that any system of mathematics large enough to contain a theory of multiplication would be either inconsistent or incomplete. Inconsistency would mean any statement and its negation were simultaneously true; incomplete means that there are statements such that both they and their negations are consistent with the axioms. This means that they cannot be proved, and there are even contrived statements of this kind that are self-evidently true. Many mathematicians today are quite happy with the set of axioms known as *Zermelo-Frankel* with choice, after Abraham Frankel, who improved Ernst Zermelo's original axiomatization, and the acceptance of the Axiom of Choice, but set theorists and logicians are aware that there are many other

choices that could be made. They feel uncomfortable with the idea that the world of mathematics could be one of many things and we will never know which one is true, and they reach back to the early days of non-Euclidean geometry for an example of what it is like to lose an easy sense of the truth of mathematics.

Perhaps the ultimate lesson of the discovery of non-Euclidean geometry is that it forced mathematicians to think clearly about what geometry is about, and to come up with a coherent set of abstract ideas for it. At first, they had two systems, both in need of further rigorization, Euclid's and that of Bolyai and Lobachevskii. Then, with Riemann, they had infinitely many off-the-peg accounts of metrical geometries. Then they came up with axiomatic geometries and axiomatic systems, which needed to be checked for self-consistency, and then it emerged that the logic governing their use also needed to be made precise. This forced a remarkable degree of clarity on mathematics and a comparable reduction on the reliance on naïve intuition. But it also proposed to workers in physics, and then in other fields, a clearer sense of what a rigorously formulated theory can provide, and how it can be tested by, and even guide, experiment, and to enrich and articulate a complicated relationship between abstraction and observation that is a characteristic feature of the modern world.

Jeremy Gray

See also Axiomatic Theory; Geometry, Euclidean; Mathematics, 19th Century; Mathematics, 20th Century; Physics, Gravitation Theory; Truth

Further Readings

- Braver, S. (2011). *Lobachevski illuminated*. Mathematical Association of America.
- Einstein, A. (1961). *Relativity: The special and general theory* (R.W. Lawson, Trans.). Crown Press.
- Gray, J. J. (Ed.). (2001). *Janos Bolyai, non-Euclidean geometry and the nature of space*. MIT Press.
- Gray, J. J. (2011). *Worlds out of nothing: A course on the history of geometry in the 19th century* (2nd ed.). Springer.
- Hilbert, D. (1990 [1971]). *The foundations of geometry* (L. Unger, Trans., 10th English edition of the 2nd German edition). Open Court. (Original work published 1902)

GEOPHYSICS

See Geology

GERM THEORY

The germ theory of disease is the contemporary scientific theory for causation of infectious diseases. In this theory, the central tenet is that microbial organisms (e.g., bacteria, viruses, protozoa, prions, and fungi) are the causative agents behind many diseases.

From before recorded history, infections have plagued both animals and humankind. The effects were devastating, and it was understood that when a disease appeared, social distancing or quarantine would normally cause the spread to stop. Each disease, however, brings with it a different challenge when it comes to managing and defeating an infectious outbreak. To achieve these objectives requires an understanding of the causative agent, its source, and what it takes to treat it. Throughout the centuries, and with the advent of the Enlightenment, many of these questions have been answered, at least in part. To arrive at the germ theory of disease required overcoming obstacles in understanding just what a microbe was and, until the invention of the microscope, it was nearly impossible to move from incorrect perceptions. Through the work of many scientists and physicians, we now have a working theory that has stood the test of time, with some variation, but newer ways of defining the causation of infectious diseases are ongoing.

Origins

The scientific journey to consolidate evidence that many diseases are caused by germs, or microbial organisms, has been a long one. From biblical times, where quarantine was first described in written form, evidence existed that an outside entity or force was responsible for causing disease, such as leprosy (a.k.a. Hanson's disease) or sexually transmitted diseases. The ancient Romans and Greeks also left records of infections that were contagious and matured into plagues. Between

460 and 400 B.C.E., the general and historian Thucydides documented symptoms and events surrounding the plague of Athens, which added to the death toll caused by the Peloponnesian War. At the time, though, pathogens were thought of as *seeds of disease* that caused disease. During this period, the philosopher Aristotle put forth the notion that contagious diseases were caused by spontaneous generation. Spontaneous generation suggested that organisms developed without the need for sexual reproduction and often formed from nonliving matter, such as dust. It was not understood that adult flies laid eggs on the surfaces of rotting food, which hatched into larvae (maggots). The eggs were not visible to the human eye, so it was thought that they just appeared due to some unknown force.

The idea that a common agent was causing disease began to mature during the Middle Ages after a Persian physician, Ibn Sina (Avicenna), wrote about contagious disease in his medical text, *The Canon of Medicine*. He used examples of known epidemics to teach about contagion, and the miasma theory became the accepted basis for understanding the causes of infectious diseases. The miasma theory set forth the principle that the agent responsible for causing infectious diseases was *bad air* or a form of noxious vapor that all the sick were exposed to. It was believed that a miasma could be identified by the foul odors that the infected often emitted while sick. The source of the miasma was predominantly described as being from the air, water, or food. There was no discrimination between a source that was independent of a miasma (e.g., contaminated water) and that which spread from person to person.

By the time of the Black Death, a devastating pandemic of the bubonic plague in the mid-1300s that was later understood to be caused by the spread of *Yersinia pestis*, a bacterium that is spread from fleas to rats, it became clear that there was some particle that was being spread from person to person. The information learned from the geographic spread and symptoms of the victims in this plague made it clear that direct contact was not always necessary for the spread of a contagious disease. The Italian physician Girolamo Fracastoro discussed this in his book *De Contagione et Contagiosis Morbis* in 1546.

In 1668, another Italian physician, Francesco Redi, put forth his theory to refute the cause of

infectious diseases through spontaneous generation. He employed the scientific method to perform an experiment where he placed meatloaf and egg in three separate jars; one was left open, another was sealed, and the last he covered with gauze. Within a few days, the meatloaf in the open jar was covered with maggots, while the jar covered in gauze only had maggots growing on top of the gauze. The sealed jar had no sign of maggots growing. Within that period, several more experimental results were offered to refute the validity of spontaneous generation, including those of Pier Antonio Micheli, John Needham, and Lazzaro Spallanzi, which, taken together, showed that living matter could not possibly arise spontaneously from nonliving matter. There had to be more to the appearance of microbes than random events.

More supportive information came forth during the Enlightenment (17th and 18th centuries), when the scientific method was diligently employed to answer many questions about the workings of the physical and organic world. Perhaps most influential at that time was the Dutch microscopist Antonie von Leeuwenhoek, who observed yeast cells undergoing division on a slide with foam from beer that had been fermenting. He also observed that fermentation was not possible in sterile air or with pure oxygen, which led him to realize that airborne microorganisms were responsible for the initiation of fermentation, not spontaneous generation.

Later, in 1859, a French scientist, Louis Pasteur, would suggest that microorganisms (germs) present in the body were the cause for infectious diseases and laid to rest the theory of spontaneous generation. The experiment involved boiling meat in a swan's-neck flask, which prevented particles from falling into the broth while still allowing air to enter the chamber. There was initially no growth in the flask, until the flask was turned in a fashion to allow particles to enter the chamber. Soon after the position changed, the broth became cloudy. The idea of spontaneous generation was nonetheless difficult to uproot, even with many experiments disqualifying it as a sound explanation of how microbes (and maggots) appear in different media (e.g., meat, broth).

In 1864, Joseph Lister, a professor of surgery at Glasgow University, learned of Pasteur's germ theory and determined to apply its principles to the problem of surgical infections. The cause of

surgical infections was not well understood, and patients who underwent surgical procedures often died from sepsis or gangrene. Many new innovations in the performance of surgery helped a patient endure a procedure, such as the invention of anesthetics, but infections remained a formidable obstacle to the uncomplicated healing of the postsurgical patient.

In his personal time, Lister, with the help of his wife, used his home laboratory to explore the nature of surgical infections. He introduced his hypotheses for the causes of surgical infections in clinical trials at the hospital where he trained as a surgeon. He explored ways to prevent bacteria from entering a surgical wound, using a chemical called *carbolic acid*, which he later called *an antiseptic*. In suitably dilute form, carbolic acid killed the germs on contact but did not do harm to human skin, which made it a success. Lister continued to develop antiseptic techniques during and after surgery, cleaning surgical instruments he invented, and asked his staff to use weakened carbolic acid in their hand washes before working in the operating arena. He also sprayed carbolic acid in the air around the patient to kill any airborne bacteria. His ultimate test was using his new techniques to perform surgery on Queen Victoria. The incidence of surgical infections plummeted after surgeons began implementing his new techniques. Even then germ theory was slow to be picked up by pockets of scientists who still endorsed spontaneous generation.

After Lister had provided a large body of evidence that supported the germ theory, a German physician named Robert Koch expanded on this work. He published a paper in 1876 that demonstrated that a bacteria named *Bacillus anthracis* was the causative pathogen for anthrax. A few years later, he followed that discovery with showing that *Mycobacterium tuberculosis* was the causative agent of tuberculosis and *Vibrio cholerae* caused the disease cholera. In the case of cholera, Lister developed public health measures to control outbreaks of the disease. Through his experiments, he devised a list of criteria that seeks to prove causation of an infectious disease, since called *Koch's postulates*. These criteria include: (1) A suspected microorganism must be present in a diseased individual, not a healthy one; (2) the microbe has to be cultured from a diseased individual; (3) inoculation of the isolated microbe

must cause disease in a healthy individual; and (4) the same microbe must be isolated from the inoculated, diseased individual. Additionally, Koch introduced using dry heat and steam sterilization to the arsenal of methods to kill harmful microbes.

Germ Theory Denialism

During the period when Pasteur's germ theory began to gain traction, another scientist, Antoine Bechamp, a chemist and pharmacologist, put forth his theory of the cause of infectious diseases, *pleomorphism*. Pleomorphism stated that a microorganism goes through various stages during its life cycle and can develop into a bacteria, virus, yeast, or fungi, directed by environmental forces. This theory was also supported by Günther Enderlein, a German zoologist and physician. Ten years of experimentation led Bechamp to conclude that molecular granulations he called *microzymas*, which are present on the cells of plants and animals, were responsible for manifestation of infectious diseases. What differs between germ theory and pleomorphism was that germ theory focused on disease being caused by external factors, while pleomorphism pointed to the internal environment. Pleomorphism has since been disqualified as a plausible theory.

Koch's Postulates Today

The underlying tenets of Koch's postulates are still widely used today. However, after the advent of the field of virology, some modification to the criteria needed to be made. For one, it was later determined that there are some microorganisms that cannot grow in cell-free culture, such as viruses and obligate intracellular bacteria. Owing to this type of exception, other avenues of infection cannot be dismissed by the Koch's postulates criteria.

To address these variances in presentation, molecular biology has been brought in to determine whether a pathogen is present to cause a disease. For example, polymerase chain reaction is used in clinical settings for sequence-based identification of viruses and bacteria. There are now whole sequencing libraries of viruses, bacteria, and large parasites for reference. This brings to light the fact that as technology advances, so will the need to modify Koch's postulates to help determine a pathogen's role in causation. When

disease is expressed in one animal model, it does not always lead to the same disease in another animal model. For example, when a bacillus gene that causes the plague is given to *Caenorhabditis elegans*, it causes death, but when introduced to a mouse model, death does not occur. Without the knowledge that disease can be species-specific, or the newer epidemiological understanding that some species are just carriers of a disease (reservoirs) while others get sick with the disease, outcomes like the mouse model would lead to failure of Koch's postulates.

In fact, in emerging zoonoses (infections that are transmitted from animal to human), the pathway of infection can be difficult to fit into Koch's postulates. When it was observed that some viruses, such as Ebola, Sin Nombre, and Middle East Respiratory Syndrome coronavirus, do not cause disease in their natural host, but then produce a severe or deadly hyperinflammatory response in a new host, it was clear that a modified version of Koch's postulates was needed. Sometimes, infected hosts are asymptomatic, so it is difficult to apply the criteria for a diseased individual when it appears no disease is present. This is an important reason to utilize genetic sequencing to determine the presence of a virus in symptomatic and asymptomatic scenarios. Using this method is supported by contact tracing in individuals or animals where a known contact has occurred and there is a mixture of disease expression and sequence-verified presence of a virus.

To address some of these discrepancies with fulfilling Koch's postulates, newer models have been instituted that incorporate Koch's postulates, but expand in specific circumstances. One popular model, the Bradford Hill criteria, is an example of broadening the Koch's postulate tenants to not just infectious diseases but also to noninfectious disease causation. Sir Austin Bradford Hill was an English epidemiologist (scientist who studies causation of diseases using mathematical models) who developed his criteria alongside demonstrating the causal link between cigarette smoking and lung cancer in 1965. The Bradford Hill criteria encompass 10 different criteria for the determination of causation: (1) strength of association (effect size), which emphasizes that the larger the association, the more likely the association is true;

(2) reproducibility, meaning that the more experimental data that show a similar effect, the more likely it is true; (3) specificity, where a specific population exposed to a specific disease is more likely to be due to the suspected exposure factor; (4) temporality, where the exposure has to occur before disease development; (5) dose-response relationship, where the more concentrated the exposure factor is, the more severe the disease expression; (6) plausibility, where a mechanism is identified and plausible; (7) coherence, where epidemiological data are corroborated by laboratory results; (8) experiment, which is linked to reproducibility, and shows that the hypothesis of exposure factor results in disease in a specific model; (9) analogy, where there are similarities between an observed association and other known associations; and (10) when the factor is removed from a diseased individual, the disease improves or disappears. The last criterion cannot always be achieved in cases where long-term injury has occurred, but improvement would still be possible with clearing an infection.

Even though Bradford Hill criteria are still used, there are discrepancies with applying these, as well as other variations of Koch's postulates. Opponents to the use of the Bradford Hill criteria state that it cannot be used as exclusive criteria for proving causality, but the use of the criteria is helpful when applied to experimental outcome results. So, it appears that there are no perfect criteria for any given scenario, but a melding of criteria such as Koch's postulates and Bradford Hill criteria, especially during infectious disease outbreaks, is scientific common sense.

Concluding Remarks

In conclusion, the advent of germ theory produced by Louis Pasteur was an essential step to understand the pathological agent, mechanism of disease, and measures that need to be taken to control or prevent disease transmission caused by microbial organisms. As with many aspects of scientific knowledge, the Enlightenment, and application of the scientific method, enabled many early scientists to set foot on the right path to question incorrect and antiquated beliefs. Through logic and skill, and with the help of technological advances, germ theory has been expanded to incorporate the

gaps in understanding based on diseases that don't fit the set of defined criteria of many better understood infectious diseases. As technology continues to advance, and newer infectious diseases emerge, there will be even more demand to redefine causation of disease.

Mandy M. McBroom

See also Biostatistics; Cell Theory; Deduction; Experiment, Theory of; Health Care Science; Hypothesis Testing; Infectious Disease Studies; Medicine, 19th Century; Medicine, Enlightenment; Medicine, Middle Ages

Further Readings

- Eisenberg, J. N., Desai, M. A., Levy, K., Bates, S. J., Liang, S., K, Naumoff, K., & Scott, J. C. (2007). Environmental determinants of infectious disease: A framework for tracking causal links and guiding public health research. *Environmental Health Perspectives*, 115(8), 1216–1223. <https://doi.org/10.1289/ehp.9806>
- Hosainzadegan, H., Khalilov, R., & Gholizadeh, P. (2020). The necessity to revise Koch's postulates and its application to infectious and non-infectious diseases: A mini-review. *European Journal of Clinical Microbiology and Infectious Diseases*, 39(2), 215–218. <https://doi.org/10.1007/s10096-019-03681-1>
- Koopman, J. (2004). Modeling infection transmission. *Annual Review of Public Health*, 25, 303–326. <https://doi.org/10.1146/annurev.publhealth.25.102802.124353>
- Lewnard, J. A., & Reingold, A. L. (2019). Emerging challenges and opportunities in infectious disease epidemiology. *American Journal of Epidemiology*, 188(5), 873–882. <https://dx.doi.org/10.1093%2Faje%2Fkwy264>
- Ramakrishnan, L. (2020). *Mycobacterium tuberculosis* pathogenicity viewed through the lens of molecular Koch's postulates. *Current Opinion in Microbiology*, 54, 103–110. <https://doi.org/10.1016/j.mib.2020.01.011>
- Scott, E. A., Bruning, E., Nims, R. W, Rubino, J. R, & Ijaz, M. K. (2020). A 21st century view of infection control in everyday settings: Moving from the germ theory of disease to the microbial theory of health. *American Journal of Infection Control*, 48(11), 1387–1392. <https://doi.org/10.1016/j.ajic.2020.05.012>
- Sergre, J. A. What does it take to satisfy Koch's postulates two centuries later? Microbial genomics and *Propionibacteria acnes*. *Journal of Investigative Dermatology*, 9, 2141–2142.

GÖDEL'S INCOMPLETENESS THEOREMS

Gödel's incompleteness theorems belong to the so-called *limitation theorems*. They indicate some inner limitations of the axiomatic method. To this group also belong the Löwenheim-Skolem–Tarski theorem (on the existence of a countable model for any consistent theory), Tarski's theorem (on the undefinability of the concept of truth), and Church's theorem (on the undecidability of the predicate calculus). All of them have far-reaching consequences for logic, for the foundations of mathematics, and for the philosophy of mathematics. They are also important for general philosophy, in particular for epistemology and methodology. However, in the latter case, the consequences depend on various additional factors, for example, on interpretations and on whether the precise assumptions of those theorems are really fulfilled.

Historical Background of Gödel's Theorems

Ever since Plato, Aristotle, and Euclid, the axiomatic method has been considered as the best method of justifying and organizing scientific knowledge. *Elements* of Euclid (4th century B.C.E.) are treated as a pattern of this method. They established a paradigm in mathematics and a pattern of a scientific theory which have been admired and followed. Since Euclid and until the end of the 19th century, mathematics was in principle developed and presented in the form of axiomatic theories (in practice, rather quasi-axiomatic) based—according to methodological principles formulated by Aristotle—axioms, postulates, and definitions. According to this, proof was the basic method of justification in mathematics. Every theorem should be proved; a proof was in fact the sole accepted method for convincing readers of the truth of claimed statements. In research practice, however, the method was not precise enough. In fact, theories have been formulated in imprecise colloquial language, since there were no precisely formulated and defined principles and inference rules. The situation changed with the development of mathematical logic at the turn of the 20th century. This enabled the introduction of formalized axiomatic systems, which provided strict and precise frameworks for constructing

mathematical theories and, in particular, made the concept of a (mathematical) proof precise. Some difficulties, however, appeared to be connected with some limitations of the formalism of logic—indicated just by the limitation theorems.

At the end of the 19th century, the German mathematician Georg Cantor developed a new theory called *set theory*. It provided a useful concept of a set, which—together with the new mathematical logic—enabled the development of mathematics in a more strict and precise way. It also provided the theory of (actual) infinity—an unavoidable mathematical concept. However, Cantor's concept of a set was not precise enough, and soon some paradoxes appeared. Various solutions have since been proposed.

One of the proposals was the reduction of mathematics to logic. This idea forms so-called *logicism*—one of the main trends in the philosophy and methodology of mathematics founded by German philosopher and logician Gottlob Frege (1848–1925) and developed by English philosophers and logicians Bertrand Russell (1872–1970) and Alfred North Whitehead (1861–1947). Another proposed solution was intuitionism, founded by the Dutch mathematician Luitzen Egbertus Jan Brouwer (1881–1966) and developed by his student Arend Heyting (1898–1980). It treats mathematics as an activity of the human mind, based on a primitive intuition of natural numbers and not as a collection of axiomatic theories.

A third proposal was the conception of German mathematician David Hilbert (1862–1943), called *formalism*. It turned out to be most promising one. Hilbert proposed to secure and justify classical mathematics (dealing with the actual infinity) with the help of finite objects which are directly and intuitively given. He proposed to use syntactical means; that is, to refer only to formulas and abstracting (for methodological purposes only!) from their meaning. According to Hilbert, this should be done by considering mathematical theories as formalized theories based on classical logic and by proving that they are consistent and complete. Add that a theory is consistent if no pair of inconsistent sentences like *A* and *not-A* can be proved in it, and it is complete if any problem formulated in the language of the theory can be solved (decided) on the basis of its axioms. In this approach a proof (called: *formal proof*) was

simply (reduced to) a sequence of formulas beginning with axioms, such that each member of it is either an axiom or is obtained from previous members by one of accepted inference rules. Such sequences are in fact finite objects given directly and intuitively, hence they can be investigated by finite, syntactical methods.

Hopes that Hilbert's program could be realized were raised by the Completeness Theorem proved by Kurt Gödel in his doctoral dissertation *Über die Vollständigkeit des Logikkalküls* (1929). It showed that means of the classical logic are in a certain sense sufficient and rich enough, that one can hope that they suffice to reconstruct any proof appearing in real research practice of mathematics as a formal proof.

Those hopes were dispelled, however, by Gödel's incompleteness theorems, proved some months later.

Gödel's First and Second Incompleteness Theorems

In 1930, Gödel proved a theorem called today *Gödel's First Incompleteness Theorem*. It states that

Every consistent theory *T* based on an effectively given set of axioms and containing the arithmetic of natural numbers is essentially incomplete, i.e., it is incomplete and any of its consistent extensions (based on an effectively given set of axioms) is also incomplete.

Add that a theory *T* is incomplete if and only if there are sentences *A* formulated in the language of the theory *T* which cannot be decided by *T*; that is, neither *A* nor *not-A* (symbolically: $\sim A$) can be proved in *T*. Such sentences are called *undecidable* (in *T*).

The theorem was announced for the first time during the Second Conference on Epistemology of Exact Sciences held in Königsberg, East Prussia, September 5–7, 1930. Gödel presented there a 20-minute talk devoted to the results of his doctoral dissertation and a day later he took part in a discussion on the foundations of mathematics and announced his First Incompleteness Theorem. The theorem was published in January 1931 in the paper “Über formal unentscheidbare Sätze der ‘Principia Mathematica’ und verwandter Systeme. I” (*Monatshefte für Mathematik und Physik* 38

(1931), 173–198). Add that in June 1932, this paper was presented by Gödel to the University of Vienna as a *Habilitationsschrift* and on the basis of it he was granted in March 1933 the *venia legendi* and became *Privatdozent*.

At the end of the paper “Über formal unentscheidbare Sätze . . .” Gödel announced also what became known as *Gödel's Second Incompleteness Theorem*. It says that:

No consistent theory T based on an effectively given set of axioms and containing the arithmetic of natural proves its own consistency—in fact the proof of the consistency of T requires means stronger than those of T itself.

Gödel promised to publish the proof of this theorem in the second part of the paper, which would be written soon. In fact the second part was never written and Gödel never published a proof of the Second Incompleteness Theorem. He explained this by saying that in the meantime the result was promptly accepted and hence there was no need to publish it. Add that the hint to the proof given by Gödel in the indicated paper has turned out to be incorrect. The first correct proof was published by D. Hilbert and P. Bernays in the second volume of their monograph *Grundlagen der Mathematik* (1939).

The First Incompleteness Theorem was proved by a sophisticated method. Gödel used the idea of representing formulas of a theory by natural numbers (so-called *arithmetization of syntax*) introduced in fact already by Gottfried Wilhelm Leibniz (1646–1716) as well as the Liar paradox known already by the ancient Greeks. He observed that by replacing truth in this paradox by provability (which in contrast to truth can be defined—via arithmetization—in the language of the theory), one can construct a sentence in the language of the considered theory stating of itself (a so-called *self-reference sentence*) “I am not provable in the theory.” This was just the undecidable sentence (called a *Gödel sentence*) indicating the incompleteness of the theory. The same procedure can be now applied to any consistent extension of the given theory. Hence any such extension is also incomplete and consequently the theory is essentially incomplete.

The method of arithmetization of the language of a theory (of its syntax) enables the translation of

metatheoretical statements (= statements about a theory) into sentences of the language of the theory itself. Hence, a metatheoretical statement “theory T is consistent” can be translated into the language of the theory T . The Second Incompleteness Theorem shows that such a sentence is not provable in T ; more exactly, one proves that this sentence is equivalent (in T) to Gödel's undecidable sentence. However, it has turned out that the way in which the translation is done plays here a crucial role.

Some Technical Details

Among the important tools used in proofs of Gödel's theorems is the concept of effectiveness and of computable function and relation. Consider functions and relations in the set of natural numbers $\mathbb{N}$. Recall that a function f is said to be computable if and only if there exists a mechanical method by which for any arguments $a_1, \dots, a_n$ the value $f(a_1, \dots, a_n)$ can be calculated in a finite number of prescribed steps. Similarly a relations R on $\mathbb{N}$ is said to be computable if and only if there exists a mechanical method such that for any arguments $a_1, \dots, a_n$, it can be decided in a finite number of prescribed steps whether the relations R holds or not. This intuitive notion is made strict and precise by introducing the concepts of a recursive function and relation.

DEFINITION. The set of recursive functions is the smallest set of functions containing the functions $+$ (addition), $\cdot$ (multiplication), $K_<$ (characteristic function of the relations $<$) and projection functions and closed under superposition and the operation of effective minimum.

DEFINITION. A relation R is said to be recursive if and only if its characteristic function K_R is recursive.

Denote by PA the formal arithmetic of natural numbers (called *Peano arithmetic*). It is a first order formal axiomatic theory formulated in the language with the following nonlogical symbols: 0 (zero), S (successor), $+$ (addition), $\cdot$ (multiplication) and based on the following nonlogical axioms:

$$(A1) \quad S(x) = S(y) \rightarrow x = y,$$

$$(A2) \quad \sim(0 = S(x)),$$

$$(A3) \quad x + 0 = x,$$

- (A4) $x + S(y) = S(x + y)$,
 (A5) $x \cdot 0 = 0$,
 (A6) $x \cdot S(y) = x \cdot y + x$,
 (A7) $A(0) \wedge \forall x [A(x) \rightarrow A(S(x))] \rightarrow \forall x A(x)$

where the symbol $\forall$ denotes the universal quantifier “for every . . .”.

The crucial connection between recursive functions and relations on the one side and Peano arithmetic on the other is the concept of representability. Denote by $\mathbf{n}$ (for a natural number n) its name expressed in the language of PA (it is called *numeral*). It is defined in the following way: $\mathbf{n}$ is the term $S(S(\dots S(0)\dots))$ in which the symbol S appears n times.

DEFINITION. A relations R is said to be strongly representable in PA if and only if there exists a formula A of the language of PA such that for any natural numbers $k_1, \dots, k_n$:

$$\begin{aligned} R(k_1, \dots, k_n) &\text{ if and only if } \text{PA} \vdash A(\mathbf{k}_1, \dots, \mathbf{k}_n), \\ \sim R(k_1, \dots, k_n) &\text{ if and only if } \text{PA} \vdash \sim A(\mathbf{k}_1, \dots, \mathbf{k}_n). \end{aligned}$$

where the symbol $\text{PA} \vdash A$ means that A is provable in PA (PA proves A).

In a similar way one defines the concept of representability of a function in PA:

DEFINITION. A function f is representable in PA if and only if there exists a formula A of the language of PA such that for any natural numbers $k_1, \dots, k_n$:

$$\text{PA} \vdash \wedge y [A(\mathbf{k}_1, \dots, \mathbf{k}_n, y) \leftrightarrow (y = f(k_1, \dots, k_n))].$$

THEOREM. Every recursive function is representable in Peano arithmetic PA. Every recursive relation is strongly representable in Peano arithmetic PA.

Another tool used in the proof of Gödel's theorems is arithmetization (sometimes called *gödlization*). It consists of establishing a one-to-one correspondence between expressions of the language of a given theory (in our case PA, though the procedure is general and can be applied to any formal theory) and natural numbers. Using such a correspondence one can translate metamathematical statements about a theory as a formal system into statements about natural numbers. If the

theory is PA (= arithmetic of natural numbers) or its extension, then one can speak about PA in itself! What more, fundamental metamathematical relations such as, for example, being a formula, being a sentence, being a proof of a given formula, and so on, can be translated in this way into certain recursive relations, hence into relations strongly representable in Peano arithmetic. Consequently one can speak about formal system of arithmetic of natural numbers (or other formal theory containing it) and about its properties as a theory in the system itself! This is the essence of Gödel's idea.

We will not describe here the arithmetization in details—it is tedious and complicated (details can be found, for example, in the monograph by R. Murawski, *Recursive Functions and Metamathematics. Problems of Completeness and Decidability*, Gödel's Theorems. Note only that if A is a formula of the language of PA, then the symbol $\underline{A}$ will denote the natural number corresponding to the formula A under the given arithmetization. The number $\underline{A}$ is called the *Gödel number* of A .

Using arithmetization one can define certain recursive relations and functions corresponding to various metamathematical relations and functions concerning PA. In particular, one obtains in this way the relations $\text{Prf}(x, y)$ corresponding to the metamathematical relation of being a proof of a formula. This means that if $\underline{A}$ is the Gödel number of a formula A and n is a natural number then

$\text{Prf}(n, A)$ holds if and only if n is the code of a (finite) sequence of natural numbers being Gödel numbers of formulas forming a (formal) proof of A in PA,

briefly:

$\text{Prf}(n, \underline{A})$ holds if and only if n is the code of a (formal) proof of A in PA,

But the relation Prf is recursive. Hence, it is strongly representable in Peano arithmetic PA. Let Prf denotes any formula of PA strongly representing Prf in PA. Consider now the relations:

$R(x, y) \leftrightarrow x$ is (a Gödel number of) a formula B with one free variable and y is a (Gödel number of) a proof of the formula $B(\underline{b})$.

Let us explain that $B(B)$ means here the formula B in which one substituted for the variable x the numeral corresponding to the Gödel number of the formula B (hence of the formula itself).

It can be shown that the relations R is recursive, hence strongly representable in PA. Using this fact one can construct a sentence (i.e., a formula without free variables) of the language of PA stating its own unprovability. Denote it by A_G and call it a Gödel sentence. Hence, the Gödel sentence says: "I am not a theorem of PA."

Before we formulate Gödel's First Incompleteness Theorem, we need still one notion.

DEFINITION. Peano arithmetic is said to be ω -consistent if and only if for any formula $A(x)$:

if $PA \vdash A(0)$, $PA \vdash A(1)$, $PA \vdash A(2)$, $\dots$ then $PA \not\vdash \exists x \sim A(x)$,

where $\exists$ denotes the existential quantifier, there exists $\dots$.

Hence ω -consistency means that if Peano arithmetic proves that every particular number n has the property A , then PA does not prove that there exists a counterexample for A (which does not mean that PA proves "for every x , $A(x)$ "!).

Note that the concept of ω -consistency is stronger than the concept of consistency, namely if a theory T is ω -consistent then it is consistent but not *vice versa*.

THEOREM (First Incompleteness Theorem; Gödel). If Peano arithmetic PA is ω -consistent, then the Gödel sentence A_G is undecidable in PA, that is, $PA \not\vdash A_G$ and $PA \not\vdash \sim A_G$. Hence, PA is incomplete.

It can be observed that the stronger assumption that PA is ω -consistent is needed only in the proof of $PA \not\vdash \sim A_G$. For the proof of $PA \not\vdash A_G$ the assumption of consistency of PA suffices.

Observe also that the described procedure can be applied to any ω -consistent extension T of PA based on recursive set of axioms (this is needed in order to secure the recursiveness, and hence strong representability, of the relations Prf). In this way, one obtains an extension of Gödel's First Incompleteness Theorem stating that PA is not only incomplete but even essentially incomplete.

It was shown later by Rosser that by considering another self-referential sentence one can eliminate from the First Incompleteness Theorem the

assumption of ω -consistency. The sentence of Rosser (denote it by A_R) says: "if there exists a proof (in PA) of the formula A_R then there exists also a proof (with even a smaller or equal Gödel number) of the formula $\sim A_R$." Hence, it says about itself that if it is a theorem of PA, then its negation can also be proved in PA (and its proof is not longer).

THEOREM (Rosser). If Peano arithmetic PA is consistent, then the Rosser sentence A_R is undecidable in PA, that is, $PA \not\vdash A_R$ and $PA \not\vdash \sim A_R$. Hence PA is incomplete.

Observe that since one of sentences of the pair A_G and $\sim A_G$ (similarly: A_R and $\sim A_R$) is true in the standard model of arithmetic $(\mathbb{N}, 0, S, +, \cdot)$, but none of them is a theorem of PA one obtains that there are true sentences (in this case A_G and A_R) which cannot be proved in PA! A careful analysis of the structure of sentences A_G and A_R shows that both are of the form "there does not exist x such that B " where B contains only propositional connectives and bounded quantifiers, hence of the form "for every x , $\sim B$." Formulas of such a form are called *formulas of the class Π_1* . Negations of such formulas, hence formulas of the form "there exists x such that B " where B contains only propositional connectives and bounded quantifiers are called *formulas of the class Σ_1* . It can be proved that formulas of the class Σ_1 , hence existential formulas are true in the model $(\mathbb{N}, 0, S, +, \cdot)$ if and only if they are provable in PA (are theorems of PA).

To formulate Second Incompleteness Theorem observe that the condition that theory T is consistent is equivalent to the condition: not every formula is a theorem of T (hence: there exists a formula A such that A is not a theorem of T).

Let Prf be a formula of the language of PA strongly representing in PA the recursive relations Prf and let Pr be the formula $\exists x Prf(x, y)$ (it says: there exists x such that x is a proof of y , hence: y is a theorem). Consider the formula Con_{PA} of the form $\sim Pr(0 = 1)$. It says: "the formula $0 = 1$ is not a theorem of PA," hence: "PA is consistent." Since there can exist various formulas Prf strongly representing the relation Prf , so the formula Con_{PA} is not uniquely defined—it depends on the choice of Prf . If the latter (and consequently also Pr) satisfies certain conditions called *derivability conditions*, then Gödel's Second Incompleteness Theorem holds. Derivability conditions require that the formula Prf not only numerically (in the sense of representability) corresponds to

the provability but also formally reflects some basic metamathematical properties of it. Derivability conditions were formulated for the first time by D. Hilbert and P. Bernays in the second volume of their monograph *Grundlagen der Mathematik*.

THEOREM (Second Incompleteness Theorem; Gödel). $PA \text{ non } \vdash \text{Con}_{PA}$.

Observe that the described procedure can be applied to any formal theory based on a recursive set of axioms and containing Peano arithmetic. This means that no such theory proves its own consistency.

Using techniques similar to those applied in proofs of Gödel's incompleteness theorems one can show that:

- (1) The set Thm_{PA} of (Gödel numbers of) theorems of PA and the set neg_{PA} of (Gödel numbers of) negations of theorems of PA cannot be separated by a recursive set, that is, there exists no recursive set X such that $\text{Thm}_{PA} \subseteq X$ and $\text{neg}_{PA} \cap X = \emptyset$.
- (2) The set Thm_{PA} of (Gödel numbers of) theorems of PA is not recursive.

Observe that the contents of Gödel and Rosser undecidable sentences is in fact a bit artificial from the point of view of a research practice of a *normal* mathematician working in number theory or arithmetic of natural numbers. In fact they have a metamathematical and not directly mathematical contents. Therefore, there arose a problem whether there exists examples of undecidable sentences of mathematically (in particular number-theoretically) interesting contents (whatever it could mean). It was solved by J. Paris and L. Harrington, who constructed an example of an undecidable sentence of combinatorial contents (a generalization of Finite Ramsey Theorem) and by Paris and L. Kirby, who provided an example of an undecidable sentence of a purely number-theoretic contents (based on the idea of Goodstein sequence)—for details see the monograph *Recursive Functions* mentioned above.

Reception of Gödel's Incompleteness Theorems

As noted earlier, the First Incompleteness Theorem was announced by Gödel for the first time during the Second Conference on Epistemology of Exact

Sciences held in Königsberg, September 5–7, 1930. It seems that all but a very few conference participants were unaware of the meaning and importance of those results. There was no discussion after Gödel's pronouncement. In the proceedings of the conference, Gödel's name does not appear. At least two persons among those attending the conference should have had foreknowledge of Gödel's result: Hans Hahn, the supervisor of Gödel's doctoral dissertation, and Rudolf Carnap. John von Neumann seems to have been the only participant who grasped the meaning of Gödel's theorem. During the journey home from the conference, he proved a theorem on the unprovability of the consistency of arithmetic in arithmetic itself. In the meantime, Gödel developed his Second Incompleteness Theorem.

Hilbert took part in the conference in Königsberg but did not attend the discussion during which Gödel announced his incompleteness result. He learned about them from P. Bernays only in January 1931 and was at the beginning "somewhat angry" (C. Reid). The irritation and frustration passed in course of time and Hilbert tried to approach the new situation in a more constructive way.

Another leading mathematician of that time, Ernst Zermelo did not accept Gödel's results and turned out to be one of Gödel's strongest critics, claiming that there is a contradiction in the proof by Gödel. In fact, the main obstacle to their understanding each other was that of deep disagreements in their philosophical views on mathematics.

Emil Post's results from 1920–1921 (on the decision problem for formal systems) anticipated results by Gödel on the incompleteness and undecidability of systems of first-order logic. He knew that his results were, as he wrote, "fragmentary." He never published them, and he gave up the research. Only after the publication of Gödel's results did he attempt to publish them, but his paper was rejected. Post always held Gödel's theorems in high esteem and expressed the greatest admiration for them. He never sought to diminish Gödel's achievement as, for example, Paul Finsler from Zurich did. The latter claimed that he had already obtained similar results in 1926. However, a careful inspection of his paper proved that Gödel was actually the first, but Finsler did not give in.

There also appeared some people claiming that Gödel's theorems were incorrect and false, that Gödel had simply discovered a new antinomy.

Such were the opinions of Charles Perelman (paper from 1936), Marcel Barzin (paper from 1940), or Jerzy Kuczyński (paper from 1938). All such opinions were criticized, and the arguments provided by the authors turned out to be wrong.

Interesting is the reaction to Gödel's theorems of two leading philosophers: Bertrand Russell and Ludwig Wittgenstein. It seems that Wittgenstein failed to understand Gödel's results. Russell reacted to Gödel's results in an ambiguous way. It seems that he was of the same opinion as Gödel, namely that by passing to higher types one can obtain formal systems such that the undecidable propositions constructed within each system are decidable in higher systems. In an unpublished letter to Leon Henkin, he admitted that Gödel's work was of fundamental importance and that he was puzzled by it. Gödel wrote in a letter to Robinson: "Russell evidently misinterprets my results; however he does so in a very interesting manner [...]."

Philosophical Implications of Gödel's Incompleteness Theorems

The First Incompleteness Theorem indicates certain limitations of the axiomatic-deductive method. In fact, any consistent formal theory containing arithmetic of natural numbers (and based on an effectively given set of axioms) is essentially incomplete. Hence, the whole of mathematics cannot be reduced to a single axiomatic system. Even more, a single theory cannot incorporate the whole truth about natural numbers. There will always be undecidable sentences of the form $\forall xA(x)$ such that all particular instances of A , that is, sentences of the form $A(0)$, $A(1)$, $A(2)$, . . . are provable. So one cannot construct a consistent and complete system of arithmetic, and all the more so of the whole of mathematics (cf. Hilbert's program described above). Even more, the particular formalization that is used is in fact irrelevant—Gödel's incompleteness results hold for any formalization provided the condition of recursiveness of the set of axioms is satisfied. Add that the incompleteness can be removed by allowing an *infinitary* rule of inference; for example, the ω -rule which leads from infinitely many premises $A(0)$, $A(1)$, $A(2)$, . . . to the conclusion $\forall xA(x)$ (which changes the character of the classical logic) or by admitting methods of set theory (with actual

infinity). Hence, infinite methods are necessary even in investigations of finite objects (such as, e.g., natural numbers).

Gödel's theorems revealed the distinction between provability and validity (truth). Every sentence provable in a theory is true in all its models but not *vice versa*; in other words, there are true (in the intended model) sentences that are not provable. The gap between provability and truth is big: the relation of being a theorem (being provable) can be defined by a Σ_1 formula whereas the concept of truth for formulas of a given theory T is undefinable in the language of T (cf. Tarski's theorem). Hence, the concept of semantical validity (truth) cannot be adequately characterized by syntactic validity (provability). One can say that the concept of truth transcends the concept of a (formal) proof, that syntax is weaker and poorer than semantics.

Another consequence of the First Incompleteness Theorem is that methods applied and admitted in mathematics as well as the creative initiative of mathematicians cannot be *a priori* bounded or limited by formal theories. Formalized theories are not fully adequate with respect to the usual research practice of mathematicians; in particular, the concept of a formal proof is not fully adequate with respect of the concept of proof used in mathematical practice. It is not a strict counterpart of informal proofs provided by mathematicians. The latter are rather argumentations that should convince readers of the correctness and truth of claimed statement; they have rather psychological (and not strictly logical) character. Any correct methods can be used in them (and formal theories limit such methods).

Using the First Incompleteness Theorem and Gödel's Completeness Theorem, one can show that there exist (beside the standard model) a rich variety of different (nonisomorphic) models of PA (the same holds for any extension of it).

The Second Incompleteness Theorem shows that in order to prove the consistency of a theory one needs means stronger than those available in the theory itself. In particular, it is impossible to prove the consistency of classical mathematics (which uses the concept of an actual infinity) by finite methods. This struck a blow to Hilbert's program; however, it did not reject it. In fact Hilbert's program has been developed in two directions: generalized Hilbert's program (the scope of

available methods is extended to, e.g., constructive ones) and relativized Hilbert's program (for which fragments of classical mathematics the original Hilbert's program can be realized).

There exist no absolute consistency proofs—only proofs of the relative consistency (if T_1 is consistent, then T_2 is also consistent) can be provided. Mathematicians can *never* be *sure* that their theories are consistent—one can only *believe* in the consistency of the considered theory (because so far no inconsistency was discovered in it and when in the past an inconsistency has been found in mathematics then it was always possible to eliminate it).

Gödel's theorems are also connected with problems of artificial intelligence. One can ask whether the human mind functions like a machine (computer), and whether the scientific activity of mathematicians can be replaced by a computer. If one assumes that a computer works like a formal axiomatic system (and this assumption is reasonable), then, it might seem, the answer should be *no*. Gödel's theorems showing certain cognitive limits of the axiomatic method indicate also some limits of the abilities of computers. But do they indicate also limitations of the cognitive abilities of human beings? We are able to decide sentences that are undecidable in some given axiomatic theories by admitting appropriate stronger methods (axioms or inference rules). Hence, again it might seem, the answer should be in fact negative.

However computers, sufficiently augmented with new rules, can do the same. Thus it is not clear that there are any questions that humans can answer that are outside the scope of what some computer, and hence some formal system, can decide.

Roman Murawski

See also Axiomatic Theory; Church-Turing Thesis; Completeness; Intuitionism in Logic and Mathematics; Logic, Formal and Informal

Further Readings

- Gödel, K. (1986). *Collected works* (Vol. I, S. Feferman et al., Ed.). Oxford University Press.
- Hoffmann, D. W. (2013). *Die Gödel'schen Unvollständigkeitssätze. Eine geführte Reise durch Kurt Gödels historischen Beweis* [Gödel's incompleteness theorems. A guided tour through Kurt Gödel's historical proofs]. Springer Spectrum.
- Hofstadter, D. R. (1979). *Gödel, Escher, Bach: An eternal golden braid*. Basic Books.
- Murawski, R. (1999). *Recursive functions and metamathematics. Problems of completeness and decidability, Gödel's theorems*. Kluwer Academic Publishers.
- Nagel, E., & Newman, J. R. (1958). *Gödel's proof*. Routledge & Kegan Paul, Ltd.
- Smith, P. (2007). *An introduction to Gödel's theorems*. Cambridge University Press.



HEALTH CARE SCIENCE

Health care science is an interdisciplinary field of applied sciences that explores how technology, science, communication, mathematics, and engineering shape, affect, and influence health care delivery. Health care science encompasses:

- The use of theory in the sciences and humanities to advance health care so as to improve health outcomes and quality of life for patients and communities;
- Models of health care delivery such as the biomedical and biopsychosocial models, as well as patient-centered communication and culturally competent communication models;
- Social and cultural determinants of health;
- The role of eHealth and technology in improving patient health outcomes and relationships with health care providers;
- Medical and social sciences that health care professionals use to advance health care;
- Scientists working on finding new treatments and cures for a variety of health issues; and
- Health care at local, national, and global levels.

The main goal of health care science research and practice, according to Hilla Brink, Christa Van der Walt, and Gisela Van Rensburg, is to provide strong evidence upon which the improvement of health care delivery can be based.

Models of Health Care Delivery

At the heart of health care science are the different types of health care delivery and how those types have evolved historically. Health can be socially constructed and understood in numerous ways, but two frameworks have been dominant and have typically shaped health care delivery: the *biomedical perspective* and the *biopsychosocial perspective*. The paternalistic and physician-centered biomedical perspective has existed since the mid-19th century and is a derivative of Louis Pasteur's germ theory of disease. Devised by medical scientists to study and cure diseases, it is a scientific model that emphasizes health as the absence of illness. As the historically predominant model of health care delivery, it is a very linear, objective perspective that places power in the hands of the physician; it is very much focused on physical health, the concept of disease, and is interested in illnesses and experiences that are verifiable. The biomedical model has both its advantages and disadvantages. Advantages, for example, include its successes in eliminating infectious diseases in the mid-19th century and dramatically increasing life expectancy in the 20th century. Disadvantages, however, include its reductionist and exclusionary nature. Many critics have noted that the biomedical model of health care delivery is exclusionary because it privileges biological determinants of illness and does not take into consideration illness symptoms; the social, behavioral, and psychological dimensions of illness; and patients' lived experiences of their health and illness.

The paradigm shift to the biopsychosocial model began in 1977 when George Engel, a professor of psychiatry and medicine, argued for a new medical model that views and treats illness not just as scientific issues, but instead as human experiences situated within sociocultural contexts. In Engel's seminal article on the biopsychosocial model, he detailed how the biopsychosocial model allows for five main components that were excluded from the biomedical model:

1. Room for variability in how various illnesses and diseases are clinically expressed and personally expressed (e.g., the clinical definition of cancer and personal definitions of cancer);
2. The ability to understand how to relate biomedical illnesses and diseases to behavioral and psychological expressions of the illness (e.g., taking into consideration both scientific and physiological aspects of an illness, as well as patients' expressions of how the illness or disease affects them);
3. Room for more discussion of when a patient identifies as *sick* and accepts the role of patient, meaning discussion of both the biochemical aspects and psychological/social aspects of disease;
4. An emphasis on treating both the biological and mental well-being of patients; and
5. A focus on improving the patient-provider relationship.

Thus, the biopsychosocial model extends the biomedical model because it adds additional layers to the illness experience and places more of an emphasis on the patient. It also requires health care providers to sharpen their care repertoire by including medical, biological, social, and psychological skill sets to gain a more holistic approach to treating patients. Since it is more subjective and views health more holistically, it focuses on every aspect of the patient's health, that is, the mental, physical, cultural, spiritual, and emotional aspects that together contribute to a person's overall well-being.

Located within these two paradigms of health care delivery are *the voice of medicine* and *the voice of the lifeworld*. Elliot Mishler made sense of the difference between these two perspectives by noting that these two voices often exist, push,

and pull at each other within medical encounters. Although patients and health care providers might be discussing the same illness, disease, and physiological symptoms, what is truly occurring is perhaps a fissure between the *voice of medicine*, the physician's biomedical approach to exploring illness and disease, and the *voice of the lifeworld*, which includes familial influences, peer influences, and other sociocultural understandings of health, disease, and illness. For example, if a health care provider and patient are discussing the patient's diabetes, the physician might discuss the patient's diabetes in terms of A1C levels and a need for more exercise and better diet choices; from within the domain of the lifeworld, the patient may be thinking of diabetes in terms of how it will affect their family members and the ability to complete job tasks at work.

Thus, the biomedical and biopsychosocial paradigms, as well as *the voice of medicine* and the *voice of the lifeworld*, have laid the foundation for a new paradigm of health care delivery that has gained traction in both theoretical and practical circles: patient-centered communication.

Patient-Centered Communication and Culturally Competent Communication Frameworks

Patient-centered care has recently advanced to the forefront of conversations about quality care. Argued to improve patient health outcomes and satisfaction with care, patient-centered care aims to enhance the quality of personal, professional, and organizational relationships. Ronald M. Epstein and Richard L. Street posit that as it advocates for helping patients become more active with their health care by becoming more active and outspoken in the health care encounter, patient-centered care challenges centuries of physician-centered and physician-dominated models of health care. Patient-centered care has three core attributes: the consideration of patients' needs, experiences, and perspectives; the provision of opportunities for patients to interact with physicians and participate in their care; and the enhancement and strengthening of the patient-provider relationship. As part of the patient-centered care initiative, patient-centered communication was developed to improve patient satisfaction, incite

better health outcomes, and overall improve patients' quality of care.

Patient-centered communication, according to Epstein and Street, can be defined in terms of patient-provider interaction processes and outcomes. The four main components of this definition include

1. Understanding patients within their own social and psychological contexts;
2. Taking time to elicit, understand, validate, and follow up on patients' perspectives, which includes their feelings, fears, concerns, questions, and expectations;
3. Reaching a shared understanding of the patients' problems and treatments; and
4. The physician sharing the power in the interaction by allowing the patient to participate and share their preferences for care.

For patient-centered communication to work, however, it must be conceptualized as a method of care that extends past the single health care encounter to enhance the patient's satisfaction with care and overall well-being even after the interaction. Behaviors that the physician can use during the encounter to practice patient-centered communication include maintaining eye contact with the patient; nodding to indicate that they understand the patient's communication; encouraging patient participation via both verbal and nonverbal means; encouraging the patient to share their feelings about their health; providing enough information so that the patient feels confident in their understanding of the health topic at hand; asking the patient about how their family members and work context contribute to their understanding of their health; and offering encouragement and nonverbal support—smiling, nodding, or a small tap on the shoulder.

In addition to the three core components of patient-centered communication, there are six core functions of patient-centered communication that, when utilized, can contribute to better health outcomes for patients who are dealing with a variety of health issues: fostering the patient-provider relationship, exchanging information, managing uncertainty, responding to emotions, making decisions, and enabling patient self-management. Epstein and Street note that these functions

overlap and interact with each other to ultimately increase patient satisfaction and improve health outcomes. All play important roles in the same health care interaction. If, for example, a pregnant patient sees her doctor because she wants more information on the amniocentesis procedure, a potentially life-altering procedure due to its invasive nature and the nature of the test information, the successful existence of and performance of all six patient-centered communication functions could be important for her during the medical encounter.

The first function of patient-centered communication, *fostering the patient-provider relationship*, is characterized by relationships between health care providers, patients, and family members that are trustworthy, mutually understanding, and full of rapport. This function takes the relationship between clinicians and patients beyond an advisor-advisee relationship toward a mutually beneficial and equal partnership where patients and their family members feel that they are welcome to participate in the medical encounter. Participation on behalf of both the physician and patient (as well as their family members, if they are present) has the potential to make the patient-provider relationship a healing relationship that is more than an information exchange based upon diagnoses and treatment plans; rather, in addition to information, there is a connection and an exchange of social support, emotional support, advice, and trust.

The second function of patient-centered communication, *exchanging information*, serves two functions: it helps patients gain more information about the health issue at hand, and it also strengthens the patient-provider relationship because the information can reassure the patient, help the patient make more sense of their illness so that a decision can be made about treatment plans, and even reduce patient uncertainty. Information exchange can occur about a variety of topics, including but not limited to illness and disease prevention, illness and disease diagnosis, and illness and disease treatment.

The third function of patient-centered communication, *responding to emotions*, is characterized by a patient-provider relationship in which the physician listens to patients' emotions about their health, takes them seriously, legitimates their

emotions, and then responds with empathy, support, and reassurance. Given that patients' emotions about their health can differ with intensity and at different stages during a disease or illness, it is important for physicians to respond positively and reassuringly to patients' emotions and offer strategies for emotion management and patient well-being.

The fourth patient-centered communication function, *managing uncertainty*, is closely related to responding to emotions and is characterized by a physician's attempt to help their patients alleviate their uncertainty about an illness or disease. Considering that uncertainty about one's health can have detrimental physical and mental effects such as emotional distress, physicians can be patient-centered by acknowledging their patients' uncertainty, offering more information that could help patients understand their health issue more clearly, and responding with empathy so that patients feel more comfortable managing their uncertainty.

The fifth patient-centered communication function, *making decisions*, is characterized by physicians and patients working together to ensure that a high-quality decision has been made about prevention and/or treatment. Epstein and Street note that a high-quality decision is one that is based on the patient's personal values, as well as their understanding of the rationale for the decision. For example, if a pregnant patient discusses the amniocentesis procedure with her physician and is unsure about whether or not to undergo the procedure, a high-quality decision in this case would be one that does not go against her values about life, bodily autonomy, and disability; moreover, it would be a decision that is based on her understanding of what the amniocentesis procedure tests for, how it operates, and whether or not she wants the amniocentesis test information. Regardless of who makes the decision (the physician, the patient, and/or family members), Epstein and Street posit that it is important that the decision-making process is characterized by both the physician and patient participating in the process, meaning they both share information, values, and feelings about the health issue at hand.

The sixth and final patient-centered communication function, *enabling patient self-management*, occurs when physicians ensure that patients are

able to follow through with their own care after they have left the physician's office or medical consultation room and feel empowered to do so. Epstein and Street note that this function is separate because it combines previous functions (information exchange and decision-making) with enablement, which is patients' perceived ability to follow through with their treatment regimen, search for more health-related information, and seek out their physician if necessary. Overall, health care providers offer support to their patients, encourage them, and provide them with information and reassurance so that they feel empowered to manage their health.

Thus, in combination these six functions—fostering the patient-provider relationship, exchanging information, responding to emotions, managing uncertainty, making decisions, and enabling patient self-management—form the basis of patient-centered communication. Other theories and models, such as culturally competent communication, also work to eliminate misunderstandings within patient-provider encounters and overall improve health care and health outcomes.

Culturally competent communication is a model that works to reduce health disparities, improve communication between health care providers and patients, and ultimately improve health outcomes. Richard L. Street and Cayla R. Teal created a culturally competent communication model that health care providers can use to improve their communication skills and models of care when interacting with patients of a variety of cultural, racial, sexual, ethnic, and demographic backgrounds. Their culturally competent communication model includes four critical components—communication repertoire, situational awareness, adaptability, and knowledge about core cultural issues—as well as five critical communication skills, which comprise a skill set that health care providers can use to engage with patients of different backgrounds. Street and Teal posit that this model was designed to elevate health care provider sensitivity to cultural difference and sharpen and refine health care providers' communication competence.

The first component, *communication repertoire*, consists of communication skills and competencies that the health care provider should have in their *communication toolbox*. A health

care provider's communication repertoire should include competencies about how physicians should obtain information, what sorts of information should be obtained within the medical encounter, and how to negotiate with the patient if cultural beliefs or values impede decision-making.

The second component, *situational and self-awareness*, deals with ensuring that health care providers have the power of perception, or the ability to attend to patients' cues and emotions. This power of perception is twofold, though, as it also entails health care providers having the ability to recognize their own implicit perceptions and biases that may affect how they provide care to their patients. If a health care provider has both situational and self-awareness, they will be able to pick up on conversational and cultural nuances and react in a patient-centered manner.

This leads to the third component of the culturally competent communication model, which is adaptability. If health care providers have adaptability, they should be able to adapt to various patients who have different belief systems. The notion of reflection applies to all components of the culturally competent communication model, but more specifically to the adaptability component because adaptability requires that health care providers can reflect upon their patients' actions and cultural belief systems within the actual encounter. Within this model, health care providers should be adaptable and reflective enough to assess their beliefs and their patients' beliefs within the moment to ensure that the most optimal care is being provided.

The fourth component of the culturally competent communication model is *knowledge about core cultural issues*, which suggests that health care providers should have knowledge about cultural groups and social determinants of health, as well as knowledge of how their own implicit biases and perceptions can influence their care and be counteracted. Teaching core cultural issues also includes a module whereby health care providers are reminded that patients might identify with various cultural groups, yet they are also individuals with individual beliefs and values. Health care providers should approach their patients as individuals first and then assess the salience of the patients' cultural beliefs and values. For example, instead of assuming that a patient will reject a

certain procedure or treatment regimen because they are a members of a particular religious group, a physician should approach the patient and then assess how salient the patient's religious beliefs are and how much their religious beliefs will influence their health decisions. This could help avoid any stereotyping and potential embarrassment and hostile feelings in the long run.

Thus, the culturally competent communication model takes patient-centered communication models one step further by incorporating knowledge about core cultural issues as a core component of improving patient-provider relationships and achieving better health outcomes. In addition to the importance of culture, two other demographic categories that have gained traction within health care sciences are race and gender.

Health Disparities and Minority Health

In addition to focusing on different models of health care delivery and different ways to improve health outcomes, another component of health care science includes the study of minority population health, more specifically from race and gender perspectives, and how to reduce and eliminate health disparities that various minorities typically suffer from. As Thomas LaVeist notes, the study of minority health is an important component of health care science because of the dramatic increase of minority races and ethnicities in the United States over the last few decades, particularly Latino and Asian populations. Moreover, Thomas LaVeist notes that current projections indicate that, in combination, minorities in the United States will become the majority population by 2050. Thus, it is imperative that health care professionals dedicate more time and effort to explore and systematically study and understand minority health because *minority health* will soon constitute the greater share of nation's population within the next few decades.

Another reason that the study of minority health has gained traction over the last few decades is because of the increase of health disparities in the United States. LaVeist posits that in general, the distribution of ill health in the United States is set up in such a manner that racial/ethnic minorities have worse health than Whites in most contexts and with a variety of health issues. There

are considerable differences in the incidence, prevalence, morbidity, and mortality for racial/ethnic minorities in the United States as compared with the White population. National data suggest that there are persistent mortality rate disparities for racial/ethnic minorities as compared to Whites, considering that Whites have a lower mortality rate than Native Americans, Hispanic/Latinos, and African Americans. Moreover, although infant mortality rates are steadily declining in the United States, there is still a considerable gap between infant mortality rates for Whites and for racial/ethnic minorities. Thus, health care disparities exist for racial/ethnic minorities regarding a variety of health care issues due to the differences in the access, use, quality of health care, and outcomes of health care services that minorities receive. LaVeist specifies that health care disparities are *racial/ethnic* differences in the quality of health care that are rooted within injustice within the health care system, whereas health care dissimilarities are racial/ethnic differences in health care that are not caused by inequities and injustices.

A myriad of theories can be used to understand and explain health disparities and can be grouped within three categories: socioenvironmental, psychosocial or behavioral, and biophysiological. *Socioenvironmental theories* such as risk of exposure theory, resource deprivation theory, and racial/ethnic segregation theory postulate that health disparities exist because of racial/ethnic differences in the communities in which people live. Different community factors that could influence a person's health include environmental issues such as toxins or pollution, crime, and lack of access to health care and healthy food, to name a few. For example, if an African American or Latino person lives within a rural neighborhood, they might not have access to a grocery store that serves fresh, healthy food or access to a health care clinic. These factors could serve as serious detriments to their care. The second category, the *psychosocial or behavioral category*, includes the weathering hypothesis, John Henryism, and immigration and acculturation theories; LaVeist notes that these theories are related to the aspects of a person's culture that influence their health behaviors. In terms of immigration and acculturation theories, for example, many patients who emigrate

from other countries might be reluctant to accept Westernized views of medicine and Westernized medical practices at the expense of their culturally related health beliefs or practices. In certain cases, this could lead to a clash of disparate cultural views of medicine and result in adverse or negative health outcomes. For example, if a pregnant woman emigrates to the United States from a rural area in Mexico and her American physician recommends she undergo a battery of prenatal screenings and testing, she might be reluctant to accept the tests because perhaps they are not available or highly encouraged in her hometown, they might go against her cultural or religious beliefs, or she might not completely understand how the tests operate and what they test for. The third category, *biophysiological theories*, includes theories such as the slavery hypertension hypothesis and consists of theories that explain how and why there might be certain genetic and biological differences that affect certain racial/ethnic groups.

Thus, health care disparities for racial/ethnic minorities in the United States exist across a wide range of illnesses and diseases and are caused by factors such as low socioeconomic status, lack of access to adequate health care, low health literacy, and racial/ethnic injustices, to name a few. Another arena in which health care disparities exist is women's health, which has also gained traction in health care science research.

Women's Health Care

Along with the study of minority health and health disparities, women's health care has also recently moved to the forefront of health care science. Health care scholars took notice of the health disparities that occurred between men and women and created the women's health care subfield, which focuses specifically on health promotion, mental and emotional wellness, sexual and relational wellness, and chronic illnesses and diseases. Given the reality that men and women's bodies are indeed different, and grounded in the belief that women deserve health care that focuses on their bodies, the women's health care subfield educates women and empowers them to learn about their bodies and also make women's health care a priority at local, national, and international levels. Moreover, women's health care is rooted in

feminist ideals of equality and justice for all women. For example, Cheryl Kolander, Danny Ramsey Ballard, and Cynthia Kay Chandler note that women's health in a global society should become a global priority because, across the globe, women share similar health concerns such as poverty, discrimination, violence, limited access to health-related information, lack of reproductive rights, limited child care services, women-specific disorders and diseases, and gynecological health issues. Thus, the women's health care subfield works to eliminate sexism, misogyny, and injustice in women's health and seeks to improve health care outcomes for all women.

In addition to being rooted within feminist ideals, the women's health care subfield is also rooted within second-wave feminism's women's health movement in the United States. Before the 1960s, women had limited influence in the women's health arena. Angry about the inferior health care that they were receiving and disillusioned with the status of women, women's oppression, and the construction of women's bodies within the health care arena, women collaborated and created the women's health movement. Created at the same time as the civil rights movement, the women's health movement, comprised of feminist and women's health scholars, sought to create equal rights for women within the health care arena. In 1969, 12 women met during a women's liberation conference in Boston and engaged in feminist consciousness raising about their dissatisfaction with their health care physicians. Deciding to pool together their knowledge, The Boston Women's Health Collective was formed and created the landmark text *Our Bodies, Ourselves* in 1970. This text contained information about women's health issues and is still available today. Many feminist scholars and women's health scholars consider this text to signify the beginning of the women's health movement. In addition to exploring women's health from feminist perspectives, the women's health care subfield also explores women's health from a gender disparities perspective. Cheryl Kolander and colleagues note that poverty not only disproportionately affects racial and ethnic minorities but also disproportionally affects women. Moreover, women in the United States are also more likely to report suffering from serious mental illness and other health issues.

Additionally, women's health care research has moved to the forefront of health care science because of the ethnic, class, and gender bias in health research. Historically, a majority of scientific research studies in health care science examined White, middle-class men's health and then extrapolated that data and applied it to women's bodies and racial/ethnic minorities' bodies. This approach has obvious shortcomings, namely that it excluded women and racial/ethnic minorities from important research and could have potentially led to negative health outcomes. This gender and ethnic bias ignored the ways in which gender and ethnicity contribute to health issues and health outcomes and also tried to apply White, middle-class men's health findings to other demographic populations. Thus, the women's health care subfield seeks to understand the unique health care issues that women face and improve health care access and outcomes for women.

Conclusion

Health care science is an interdisciplinary and transdisciplinary field that uses medical science and social science approaches to explore health care delivery and improve health care outcomes. By embracing studies and perspectives from technology, communication, mathematics, and engineering, health care science explores topics such as models of health care interaction, patient-centered communication and culturally competent communication, racial/ethnic health disparities, women's health, and class, gender, and race disparities within health care.

Leandra Hinojosa Hernandez

See also Biopsychosocial Model; Complementary Medicine; Infectious Disease Studies; Medicine, Contemporary

Further Readings

- Brink, H., Van der, W. C., & Van Rensberg, G. H. (2012). *Fundamentals of research methodology for healthcare professionals*. Juta and Company Ltd.
- Engel, G. (1977). The need for a new medical model: A challenge for biomedicine. *Science*, 19(4286), 129–136. <https://doi.org/10.1126/science.847460>

- Epstein, R. M., & Street, R. L. (2007). *Patient-centered communication in cancer care: Promoting healing and reducing suffering*. National Cancer Institute, National Institutes of Health.
- Epstein, R. M. & Street, R. L., Jr. (2011). The values and value of patient-centered care. *Annals of Family Medicine*, 9(2), 100–103. <https://dx.doi.org/10.1370%2Fafm.1239>
- Graham, N. O. (1995). *Quality in health care: Theory, application, and evolution*. Aspen Publishers.
- Kolander, C. A. (2011). *Contemporary women's health: Issues for today and the future*. McGraw-Hill.
- LaVeist, T. A. (2005). *Minority populations and health: An introduction to health disparities in the United States*. Jossey-Bass.
- Mishler, E. G. (1984). *The discourse of medicine: Dialectics of medical interviews*. Ablex.
- Sahler, O. J. Z., & Carr, J. E. (Eds). (2012). *The behavioral sciences and health care*. Hogrefe.
- Teal, C. R., & Street, R. L. (2009). Critical elements of culturally competent communication in the medical encounter: A review and model. *Social Science & Medicine*, 68(3), 533–543. <https://doi.org/10.1016/j.socscimed.2008.10.015>

HYPOTHESIS TESTING

Hypothesis testing constitutes a quantitative method for refuting or supporting theories based on observations from a sample. This entry provides an overview of a commonly used statistical procedure that combines Ronald Fisher's statistical theory with Jerzy Neyman and Karl Pearson's approach to drawing inferences from a sample to the population. Historically, these two competing statistical theories produced heated debates between their originators. Later, however, ideas from the two approaches were merged in the social sciences and are often treated as complementary ways for making inferences about the probability of the existence of the hypothesized relationship in the population of interest.

Hypothesis testing in social sciences typically involves testing a hypothesis of relationship between variables (H_1) against a hypothesis of no relationship (H_0). For example, H_1 will suggest that watching educational television programming is associated with better school readiness skills in 4-year-old children, whereas H_0 will maintain that

there is no association between watching educational television programming and school readiness. Statistical tests enable researchers to reject the null hypothesis and conclude that the relationship between the variables observed in the sample was unlikely to occur by chance. This can be interpreted as evidence in support of the theory about the relationship between variables on the population level. However, hypothesis testing cannot definitively confirm or refute the theory. Inferences from test statistics are probabilistic; therefore, they are open to a certain (known) probability of error. The study might fail to detect a relationship between variables that exists in the population (e.g., in the example above—researchers conclude that educational television viewing and school readiness skills are not associated, when in fact they are). Alternatively, a relationship that *does not* exist in the population can be found in the sample, leading to the erroneous conclusion that the variables are related in the population (e.g., researchers conclude that educational television viewing and school readiness skills are positively associated, when, in fact, they are not).

Despite the widespread use of null hypothesis testing, there is an ongoing controversy concerning the validity of this approach. A number of alternatives to null hypothesis testing, such as Bayesian statistics, have been advanced. Within the frequentist statistical approach, reporting confidence intervals and effect sizes to supplement or replace hypothesis testing became a norm in many research fields.

The Logic of Hypothesis Testing

Statistical hypothesis testing rests on the principle of logic called *modus tollens* (from Latin: “the way that denies by denying”). Instead of directly examining the probability of the existence of the hypothesized relationship, this statistical approach aims to show that the opposite (lack of a relationship) is unlikely, and therefore, the hypothesized (i.e., the existence of a relationship) is highly probable. In other words, the statistical test looks at the probability of data (D) to occur in the population in which a no-relationship hypothesis (H_0) is true, $P(D|H_0)$. While researchers, from a theoretical standpoint, are typically interested in demonstrating a relationship between constructs, evidence

in support of this hypothesis is established by showing that the opposite scenario (lack of relationship) is unlikely.

Stating the Hypotheses

The first step in hypothesis testing involves stating the null (H_0) and the alternative (H_1) hypotheses. The alternative hypothesis is the theory driven hypothesis that the researchers hope their data will support. For instance:

H_0 : There will be no difference between the GRE scores of students who took a GRE preparation course and those who did not take it.

H_1 : Students who have completed a GRE preparation course will score higher on a GRE test, compared with students who did not take the course.

In this example, the null hypothesis is a nil hypothesis, as it literally states that there is zero relationship between the variables of interest (here—taking a preparation course and test scores). Researchers almost uniformly use nil null hypotheses, and most statistical packages only allow testing such hypotheses. It is, however, possible to state non-nil null hypotheses, where the null hypothesis allows for a theoretically negligible difference to exist. In the example above, such a set of hypotheses can, for instance, read:

H_0 : Students who have completed a GRE preparation course will score *no more than 10* points above students who did not take the GRE preparation course.

The alternative hypothesis will therefore state:

H_1 : Students who have completed a GRE preparation course will score *10 or more* points above students who did not take the GRE preparation course.

Some researchers argue that non-nil hypotheses are preferable to the pervasive nil null hypotheses. First, non-nil hypotheses are more accurate and realistic. Two means will essentially always be at least somewhat different, and two variables rarely will have a precisely zero correlation. Thus, in actuality, a nil null hypothesis is always false. The question should be whether there is a theoretically or practically meaningful difference between the

means rather than whether the means are different at all. A non-nil hypothesis forces researchers to incorporate practical and theoretical significance considerations in hypothesis testing. In the example above, researchers should ask themselves what constitutes a meaningful effect of a GRE preparation course. What should be the minimal consequential difference in GRE test scores between those who have completed a preparation course and those who did not take the course? For instance, does a mean difference of one point in GRE scores make the cause worthwhile? Instead of rejecting the null hypothesis due to any difference, no matter how trivial it is, theory and prior research findings can be used to inform a non-nil null hypothesis, specifying the minimal effect that has to be observed in the sample. Yet, despite the theoretical advantages of this approach, nil null hypotheses are conventionally used.

Parametric and Nonparametric Test Statistics

Hypothesis testing is based on various test statistics that compare the actual data obtained from the sample and expected distributions. Most commonly in social sciences, researchers use parametric tests that depend on assumptions about the variables' distribution in the population. These typically include assumptions of independence of observations, normality, and homoscedasticity. Violation of these assumptions might bias the estimates yielded by parametric statistics. It may be possible to reduce skewness in asymmetrically distributed data using log, square-root, and other transformations. Yet, in this process, data units lose their original meaning and therefore make the interpretation less intuitive. One evident example is a regression analysis where results based on analysis of raw data can be interpreted as increase in one unit of the independent variable leading to an increase in a certain number of units in the dependent variable. For example, based on the regression results it can be said that every hour of preparation for the test increases the score by 0.86 points. Results of a binomial regression can be interpreted to indicate that every cigarette smoked a day increases one's odds of failing the fitness challenge by 0.08. However, if the independent variables were transformed, it is no longer

possible to express the relationship between the variables in these terms. If the data were log-transformed, then the relationship between the independent and the dependent variable becomes exponential. Rather than each smoked cigarette or each hour of preparation contributing equally to the outcome, the dependent variable grows at a different rate for each unit (the first hour of preparation or the first cigarette has little effect on the exam score or the chances to pass the fitness challenge, but the 10th hour/cigarette makes a substantial contribution to the outcome).

Although the results can be interpreted as average effects of the independent variable on the outcome, from a purely theoretical standpoint transformation makes the test less appropriate for the theory, if the theory originally predicted a linear relationship between variables. In other words, while data transformation assists in meeting the analysis requirements, it also somewhat changes the hypothesis being tested.

Instead of transforming nonnormally distributed data, researchers may consider tests that are more robust to violations of assumptions. Nonparametric tests that do not rely on assumptions concerning the distribution of the examined variables include, among many others, Kruskal–Wallis analysis of variance, Mann–Whitney U (as an alternative to t -test), and nonparametric regression.

The p Value

Test statistics produce p values that indicate the statistical significance of the research result. The p value constitutes the probability to obtain the study result (or a more extreme effect) given that H_0 is true for the population from which the study sample was drawn. The smaller the p value, the more unlikely the result is attributed to chance. In other words, a smaller p value implies a higher probability that the H_0 is false. An arbitrary cut-off point of $p = 0.05$ is commonly used, such that a test statistic with an associated p value larger than that threshold is deemed a statistically non-significant result.

In view of common misinterpretations of test statistics, it is important to underscore what a p value *is not*. First, the p value *cannot* be interpreted as the probability of H_0 being correct. The

test statistic only tells researchers how likely a particular sample-based finding is given that H_0 was true in the population. In other words, it provides information regarding the likelihood of the data obtained, not the hypothesis itself.

Second, p values are sometimes confused with the magnitude of effects. Significance levels, as indicated by p values, are concerned with the probability that the effect observed in a sample is likely to occur in the population. However, the size of the p value says nothing about the size of that effect. A significance level of 0.034 does *not* indicate a more substantial research result or a stronger effect than a $p = 0.043$. As will be discussed below, p values depend to a large extent on sample size. Therefore, if a few more study participants are added to the sample without changing the means and standard deviations of the compared groups, then the same effect size will have different levels of significance. Moreover, owing to differences in sample size, a larger effect size might have smaller p value than a smaller effect. Finally, statistical significance is unrelated to social or clinical significance and the practical implications of the study results. Theory and knowledge of the subject are needed to judge the applied meaningfulness of the findings.

One-Tail Versus Two-Tail Test

A one-tail test is conceptually appropriate when researchers have a strong theoretical reason to hypothesize an effect (e.g., mean difference) in a specific direction, expecting the mean to be either greater or smaller than a certain value. If researchers hypothesize that students who took the preparation course score higher on the GRE test than those who did not take the course, H_1 can be: $\mu_{\text{with course}} - \mu_{\text{no course}} > 0$ and H_0 will read: $\mu_{\text{with course}} - \mu_{\text{no course}} \leq 0$ (those who took the course had the same scores or lower scores than those who did not take the course). The term *one-tail test* stems from the fact that test results that fall in only one of the tails of the distribution (in the example above, positive, right-hand tail) will lead to rejection of the null hypothesis.

A more rigorous test will consider differences in both directions, not just one. To reject the null hypothesis, a more extreme test statistic value is required. The hypothesized relationship is stated

as: $\mu_{\text{with course}} - \mu_{\text{no course}} \neq 0$ and the null hypothesis will include both $\mu_{\text{with course}} - \mu_{\text{no course}} < 0$ and $\mu_{\text{with course}} - \mu_{\text{no course}} > 0$. In this situation, the observation should fall within either the positive or the negative side of the distribution in order to lead to rejection of the H_0 of no difference. Figure 1 presents the differences in the critical regions for one-tailed and two-tailed hypotheses on a sample of distributions of means. The boundaries for a two-tailed test at $\alpha = 5\%$ (eliminating the extreme 2.5% on each side) are $z = 1.96$ and $z = -1.96$. For a one-tail test, the critical value for rejecting

the null hypothesis marks the boundary of the extreme 5% on one of the sides of the distribution at $z = 1.64$ or $z = -1.64$.

One-tail tests have greater statistical power than two-tail tests, giving researchers a better chance to detect effects in the population. However, it is generally less acceptable to use one-tail tests because researchers might miss an effect in the opposite direction. This, in turn, might have serious ethical and practical implications (e.g., when testing the effectiveness of a new drug). The more conservative two-tail test is recommended in many research

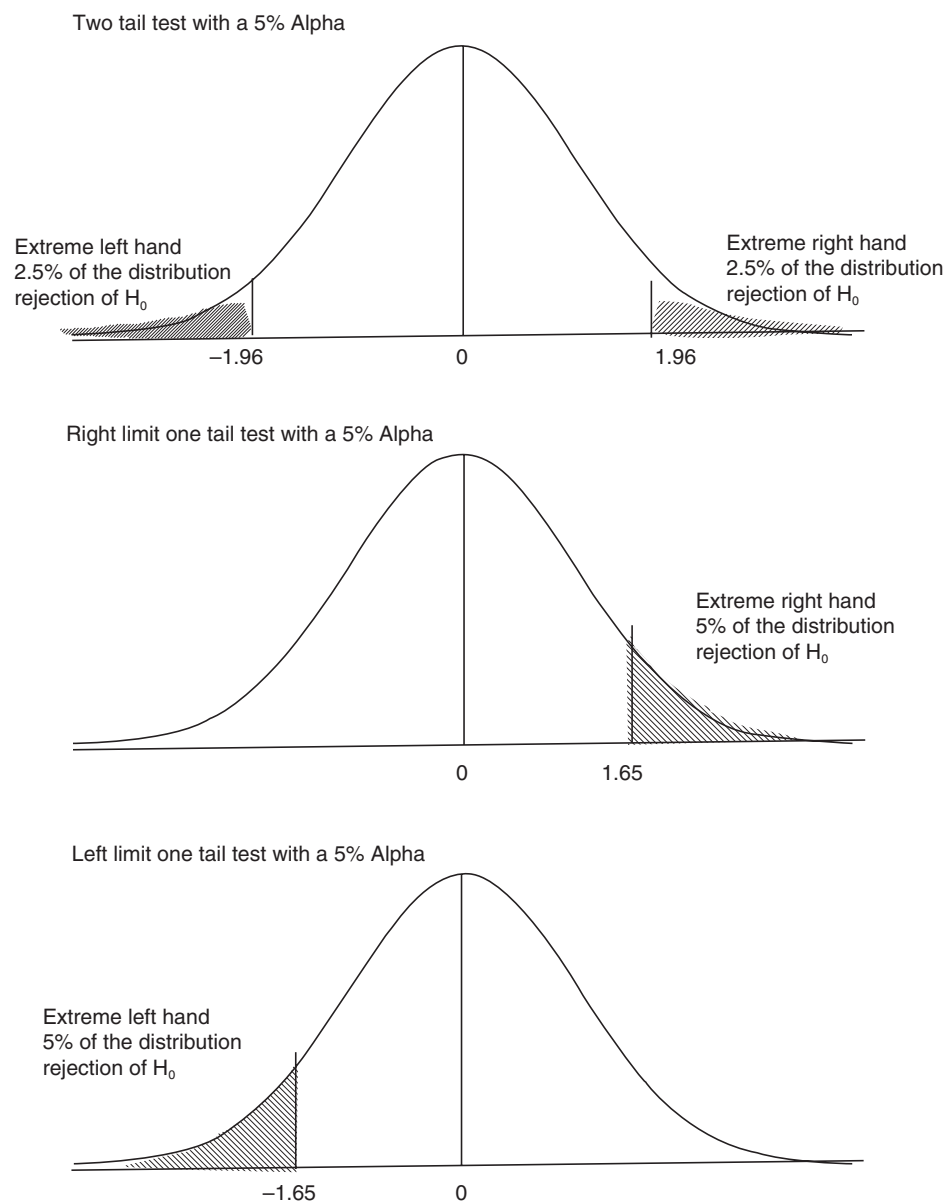


Figure 1 Areas of Rejection of H_0 in Two-Tail and One-Tail Test When $p = 0.05$.

fields even when researchers propose theory-driven, directional hypotheses. In other fields, it can be permissible to use a theory-driven one-tailed test that was proposed in advance, before the data have been collected (i.e., never after a two-tailed test yielded a nonsignificant result).

Limitations of Drawing Conclusions From Test Statistics

The probabilistic nature of hypothesis testing limits the scope of inferences that can be made on the basis of this method. Rejection of the null hypothesis does not indicate that the null hypothesis is false (the p value can never be 100%). It only implies that the null hypothesis is highly unlikely to be true. It is similarly impossible to prove the null. Failure to reject H_0 is not equivalent to accepting the null. Rather, the conclusion is that there is merely insufficient evidence for rejecting the null. In fact, hypothesis testing provides little information about the state of the null hypothesis. Hypothesis tests inform about the probability of data occurring in a population, given that H_0 is true ($P(D|H_0)$), but they do not allow one to conclude the inverse probability—the likelihood of H_0 being true, given the data observed ($P(H_0|D)$).

Bayesian statistics provide an alternative approach favored by many critics of traditional methods of hypothesis testing, because this method assesses the probability of each hypothesis more directly. Moreover, the Bayesian hypothesis testing approach considers both the observed data and its prior probability in order to determine the likelihood that the researcher's hypothesis is true, given the evidence.

Errors in Hypothesis Testing

The goal of hypothesis testing is to use data drawn from a sample to learn about the population from which the sample was obtained. However, the process of making inferences from the sample to the population is open to a number of errors. Table 1 presents the definitions of two errors that are most commonly discussed in the literature.

Since not all random samples drawn from a given population represent the population equally well, on rare occasions researchers might base their

statistical conclusions on an extreme sample where they find a significant association between variables otherwise not correlated in the population. Consequently, researchers make a *Type I error*: erroneously rejecting the null hypothesis. In other words, Type I error occurs when the hypothesized relationship was found in the sample and researchers concluded that this is true for the population, when, in fact, the relationship does not exist in the population. The probability of avoiding Type I error (α) is set in advance as a function of the p value. With the threshold for acceptable significance level set at the conventional 0.05, α is 95%. Researchers are therefore taking the risk of wrongly rejecting H_0 in an average of 5% of the tests.

The risk of Type I error increases with the number of tests conducted on the data. Since, for each significant test, there is a probability of 5% that the result was achieved by chance, repeated testing increases the probability that one or more of these tests will be significant. For instance, consider a study in which researchers examine the effectiveness of a new antidepressant medication compared with a placebo. The researchers use repeated measures of five dependent variables. The outcome variables include (1) self-report of the patient's mood, (2) a survey of a patient's family member about the patient's mood, (3) a reaction time test assessing the patient's mood, and two physiological markers of depression: (4) heart rate variability, and (5) prefrontal EEG asymmetry. These five measures are repeated four times: 2, 4, 6, and 8 weeks after the beginning of the treatment. This design produces a total of 20 posttest comparisons between the treatment and the placebo groups. With the p value set at 0.05, there is a 5% chance for researchers to obtain a significant test result by chance alone, which means that, on average, 1 in 20 comparisons between the groups will incorrectly yield a significant outcome. In the example above, suppose that the researchers do not obtain any significant effects of the drug on any outcome except one. They only detect a significant effect of the medication on self-reported mood at the 8-week follow-up. The researchers then conclude that the medication has a long-term effect on mood. However, with so many tests it is conceivable that the significant effect was a fluke rather than valid evidence of the true effect in the population, and the researchers'

conclusion is subject to Type I error. To maintain statistical rigor, various statistical procedures such as that of Bonferroni, of Scheffe, or of Tukey can be used to adjust the results to account for the inflation of α in multiple subtests. These adjustments reduce the so-called *experiment-wise error rate* by decreasing the required p value for each individual test.

The flip side of Type I error is Type II error. *Type II error* occurs when researchers retain H_0 when it is incorrect and falsely conclude that the relationship does not exist in the population. That is to say, whereas Type I error refers to false positives, Type II error involves false negatives. The probability of making Type II error is denoted by β , and $1 - \beta$ (i.e., the probability of not making a Type I error) is termed *statistical power*. As a rule of thumb, researchers feel comfortable with statistical power of 80% (i.e., $\beta = 0.20$).

Three factors play a role in determining statistical power: the probability of a Type I error, effect size, and sample size. First, it is important to note that there is a trade-off between Type I and Type II errors. Reducing power from 80% to 60% (i.e., setting β at 0.40 rather than 0.20) increases the probability of Type II error but makes Type I error less likely. The reverse is also true: the more liberal the researchers are in rejecting the null (e.g., use a p value of 0.10 instead of 0.05), the more likely they are to make a Type I error. However, they will also be less likely to perform Type II error and miss an existing effect. Since α is typically set at the conventional 0.05 or 0.010 level, practically speaking, researchers can improve their power only by adjusting the other two parameters (sample size and effect size).

The second factor is effect size. The larger the effect size (e.g., stronger correlation between variables, or larger mean difference and more homogeneous variances), the more likely one is to observe this effect in the sample. Third, it is easier to detect a significant effect in a larger sample. Even small effect sizes (i.e., minute correlations or tiny mean differences) are more likely to be statistically significant in a very large sample. Even very small correlations such as $r = 0.02$ will yield a very small p value when sample size is in tens of thousands.

The probability of making Type I and Type II errors for each given study is established by researchers based on cost-benefit considerations.

Customarily, however, Type I errors (assuming an effect exists when it does not) are considered to be more serious than Type II errors (missing an existing effect). Thus, by convention, the Type II error rate is set to be four times as large as a Type I error (i.e., 20% vs. 5%, respectively). Depending on the severity of the consequences of making a Type I error, researchers may choose to adhere to a more conservative threshold for rejecting the null. For example, repercussions of Type I error in the case of approving a new drug can be quite dire. Patients might suffer from side effects from an ineffective medication without receiving any of the promised benefits. Moreover, conceivably, more effective treatment will be withheld from these patients. Thus, in medical research, researchers may want to reduce Type I error by increasing α to 0.01 or 0.001 and thereby require stronger evidence to reject the null. This will come at a cost of increasing the chance of making a Type II error. However, the researchers might reason that if the drug is very powerful and produces a substantial effect size, they will have good statistical power to detect this effect, even with a smaller α . It is often better to overlook a small effect size (i.e., minimal advantages of a new medication) and to continue using the old, slightly less effective drug, than to adopt a potentially ineffective treatment, given the lack of long term, follow-up data.

Finally, a less commonly discussed type of incorrect inference is *Type III error* (γ), wherein the researchers correctly reject the null hypothesis but make an error concerning the direction of the effect. A small number of samples can produce an extreme statistic that displays the opposite trend of the effects compared to the population. For example, in the population, most students benefit from taking a GRE preparation course. However, in a very small number of samples from this population, the opposite will be true, and course-takers will score significantly lower than those who did not take the course. The test statistic warrants rejection of the null hypothesis (which is true, because there is, indeed, a difference between course takers and nontakers in the populations), but the researchers draw a wrong conclusion about the direction of the effect in the population (believing that the course reduces GRE scores rather than increases them). Accordingly, the probability of correctly rejecting H_0 equals $1 - \beta - \gamma$.

Table 1 Definitions of Type I and Type II Errors

	<i>Made a Correct Decision</i>	<i>Made an Incorrect Decision</i>
Researchers rejected H_0 based on a significant test result	$1 - \alpha$ (usually is 0.95) Researchers made a correct inference that there <i>is</i> a relationship in the population.	Type I error α (usually set at 0.05) False positive: There is no relationship between the variables in the population, but the researchers concluded that there is a relationship.
Researchers retained H_0 based on a nonsignificant test result	$1 - \beta$ (statistical power, usually 0.80) Researchers made a correct inference that there is <i>no</i> relationship in the population.	Type II error β (usually set at 0.20) False negative: There is a relationship between the variables in the population, but the researchers concluded that there is no relationship.

Type III error is typically attributed to outliers. Thus, not surprisingly, Monte Carlo simulations of Student t -tests reveal that the rate of Type III error is greater in studies using small samples and when comparing groups with discrepant sample sizes.

Alternatives to Hypothesis Testing and Additional Considerations

Despite an ongoing controversy concerning its effectiveness, null hypothesis testing dominates in many social scientific fields such as social psychology, education, and communication studies, and it is prevalent in other sciences such as public health and medicine. Opponents of null hypothesis testing advocate for a number of alternative or supplemental approaches. First, as noted briefly above, Bayesian statistics provides an alternative to traditional hypothesis testing that examines the probability of the theory-driven hypothesis more directly rather than through denial of the opposite. The Bayesian approach to hypothesis testing is also less sensitive to sample size; it does not depend on whether the test is planned or post hoc; and it accounts for the base rate of the phenomenon under examination. Recently, a number of statistical packages began to offer at least some support for this approach, and custom programs can supplement mainstream statistical software.

Second, within the realm of frequentist statistics, confidence intervals allow researchers to estimate the effect in the population, rather than to

test hypotheses. Whereas p values prompt dichotomous decisions about rejection or retention of H_0 , confidence intervals present a range within which the true population effect size is likely to occur with a certain probability (typically, 95%). Larger sample size narrows the interval, thereby increasing the precision of the estimate. If the interval includes the value zero, it is not statistically significant. Thus, proponents of this method maintain that confidence intervals provide the same information as p values with the additional benefit of specifying the upper and lower limits within which lies the true population effect.

Finally, some argue that the practice of hypothesis testing requires that researchers devote too much attention to the question of statistical significance rather than focusing upon effect sizes. In this view, not only does null hypothesis testing depend on an arbitrary cut-off point ($p = 0.05$), but statistical significance is too heavily dependent on sample size. A study conducted on a sample of 40 individuals can yield a nonsignificant correlation ($p = 0.055$), which is likely to be deemed nonpublishable. However, had one collected data from another dozen of research subjects without changing the correlation, the results would now cross the significance threshold (say, down to $p = 0.044$), turning the exact same finding to be statistically significant and, therefore, worthy of a publication. Furthermore, even minute and socially or practically insignificant differences become statistically significant in large enough samples (due to the

relationship between power and sample size). Taken together, tests for significance alone do not provide meaningful information about the theorized relationship between variables.

Recognizing the limitations of null hypothesis testing, many journal editors require supplementing these tests with information about effect size. In many scientific journals, it may no longer be acceptable to simply present a *t*-test indicating that there is a significant difference between two means. Rather, researchers are required to compute Cohen's *d*, eta-squared or some equivalent statistic that addresses the magnitude of the effect. Moreover, given that some researchers do not think of statistical significance as a binary decision, the 6th edition of the American Psychological Association's publication manual urges authors to indicate exact *p* values to two or three decimal places in lieu of *p* value range (for instance, $p = 0.044$ and not $p < 0.05$). In some scientific journals and certain areas of research (e.g., applied behavior analysis), null hypothesis testing is rarely used. Instead, graphic representation of the data and detailed descriptive statistics are used to communicate the research results.

Riva Tukachinsky

See also Statistical Inference, Bayesian; Statistical Inference, Frequentist; Statistics

Further Readings

- Cohen, J. (1994). The earth is round ($p < .05$). *American Psychologist*, 49(12), 997.
- Gigerenzer, G. (2004). Mindless statistics. *The Journal of Socio-Economics*, 33(5), 587–606. <https://doi.org/10.1016/j.socec.2004.09.033>
- Eidswick, J. (2012). A Bayesian alternative to null hypothesis significance testing. *Shiken Research Bulletin*, 16(1), 2–15.
- Ioannidis, J. P. A. (2005). *Why most published research findings are false*. *PLOS Medicine*, 2(8). <https://doi.org/10.1371/journal.pmed.0020124>
- Macdonald, P. (1999). Power, type I, and type III error rates of parametric and nonparametric statistical tests. *The Journal of Experimental Education*, 67(4), 367–379. <https://doi.org/10.1080/00220979909598489>
- Mittelhammer, R. C. (2013). Hypothesis testing theory. In *Mathematical statistics for economics and business* (pp. 523–607). Springer.

Wilkinson, L. (1999). Statistical methods in psychology journals: Guidelines and explanations. *American Psychologist*, 54(8), 594–604.

HYPOTHETICO-DEDUCTIVISM

Hypothetico-deductivism involves the testing of theories using observations. From a general statement or theory, a hypothesis is derived, which is then tested using observations. If the observations do not support the hypothesis, the hypothesis is rejected, and subsequently the theory is rejected. If the observations support the hypothesis, the theory is corroborated because testing of the hypothesis is seen as testing of the theory itself.

This is often explained as follows. Suppose we have a theory, *H*, and it predicts the phenomenon *O*. Through our observations, we observe the phenomenon *O* occurring. This does not prove or confirm, however, that *H* is true, because *O* can occur for reasons other than *H*. If *O* is not observed, however, we can reject the theory, *H*. Therefore, a theory can be refuted by a negative observation but cannot be proved by a positive observation. When theories are not refuted, they are merely corroborated and never proven to be true. This can be illustrated using a classic example. If we have a theory that all humans are mortal, we can refute this theory by showing through our observations at least one human who is immortal. If from our observations so far, all humans are mortal, then we have corroborated the theory.

Deduction is an integral part of hypothetico-deductivism. Deduction is a process whereby we derive particular statements from general statements. Deduction can be traced back to Euclidean geometry and Aristotelian logic. Hypothetico-deductivism was first introduced by the British scientist, philosopher, and historian of science William Whewell (1794–1866). Later, Karl Popper (1902–1994), an Austrian-British philosopher, also offered this as an alternative to the inductive methodology.

Hypothetico-deductivism is associated with Popper's falsificationism. In falsification, a statement is classified as false based on observations. He also emphasized that for a theory to be

considered science, it has to be falsifiable. Falsificationism follows a cautious realist ontology and was proposed as a way to deal with the problems of empiricism. As a deductivist, Popper maintains that we reason deductively. He put forward this logic to deal with the problems of induction. According to Popper, induction, which is based on experiences and finite observations, cannot provide a reliable theory. This is because inductive logic will never be able to provide a theory with certainty. Since induction is thus flawed, he proposed a different logic to develop theories.

Popper distinguishes between two categories of philosophers: verificationists and falsificationists. He himself obviously falls in the second category. According to falsificationists, scientific knowledge cannot be definitively proven or established with certainty but remains tentative forever. Rather than using observations to develop a theory (as in induction), falsificationism depends on observations to test theories or to falsify theories. This is an important aspect of hypothetico-deductivism. Popper sees science as a pursuit of truth about the universe, but theories can never be proven or confirmed. Popper rejected the notion of confirmation because it implies verification. According to this view, no amount of evidence can prove a theory to be true, but one instance of contrary evidence will prove that a theory is false. This point can be illustrated using a classic example. Even though we may observe many white swans, this does not prove that all swans are white or will be white; but an observation of a single black swan will immediately prove that our generalization (that all swans are white) is false.

Popper's method starts with a tentative idea, a hypothesis or a set of hypotheses that form a theory. Based on this initial idea, a set of conclusions is made and scrutinized to ensure that it helps the advancement of knowledge regarding the phenomenon being studied. In addition to this, the logic used to deduce these conclusions is also scrutinized. After that, these conclusions are tested by making observations. If the observations are not consistent with our conclusion, it shows that our *theory* is false and must be rejected. If the theory, so called, passes our test—that is, if the observations are consistent with the theory—then the theory is *corroborated* (but *not* thereby proven to be true).

Popper set forth a few requirements for his method. First, theories are considered scientific only if it is possible to falsify them; otherwise they are not considered scientific. In order to be falsified, a theory needs to make a definite claim about the world. The more clearly a theory is stated, the more falsifiable it is. This criterion is important because it differentiates scientific theories from other forms of explanation. Second, theories should always be open to criticisms and the ultimate criticism is the attempt to refute a theory. Theories which are not open to criticism are not considered scientific. Third, if a theory is refuted, it can be modified or substituted by other theories, which will be tested further. Any theory that survives refutation is only corroborated, pending further attempts to falsify the theory. Hence, theories are only accepted provisionally for as long as these theories survive refutation.

Critiques of Hypothetico-Deductivism

Even though hypothetico-deductivism is prevalent among scientists and social scientists as well, it has been criticized by some philosophers for entailing many technical issues. One of the most important criticisms of hypothetico-deductivism is that it does not discuss the source of the propositions in a deductive theory. Popper himself did not discuss this; he was skeptical of any logic of discovery and felt that the context of discovery was psychological. He maintained that the initial discovery phase does not require a logical analysis but that logical analysis is important in the theory-testing stage. He believed that every discovery involves some form of creative intuition or perhaps simply a guess.

Popper's view that there is no logic of discovery, and no need for one, has been heavily criticized. People or more specifically scientists do not develop a hypothesis at random or by mystical intuition. It is often developed through reasons and arguments. Therefore, scientists have questioned whether a theory is derived from some creative intuition or by some form of inductive reasoning. In the Enlightenment era, the British philosopher John Locke (1632–1704) argued that everything we learn about the world is from our experiences, and we are not born with innate ideas. This view has been supported by many

philosophers and practitioners (from various fields) ever since. Scientists and philosophers find it hard to accept that theories can be developed *just like that*. Most of the time, development of a hypothesis is done inductively; indeed the term *hypothetico-deductivism*, which was suggested to deal with the problem of induction, is still very much dependent on induction. Hence, it is not a solution to the problem of induction. Therefore, the question remains as to whether initial theories or strict tests of hypotheses are possible without inductive reasoning.

Another major criticism of hypothetico-deductivism is that it cannot escape induction even in the testing stage. When a theory is not falsified, its acceptance is based on induction. If the best theory is the theory which survives severe tests, that only means that corroboration of a theory is based on inductive support. Therefore, philosophers have argued that falsification does not provide a basis for believing a theory even though Popper himself emphasized that theories are only provisional and should not be accepted as the truth. Philosophers argued that our belief in a theory is often based on its past success. This means that we are making an inductive argument to corroborate that theory. Therefore, hypothetico-deductivism includes the very problem it sets out to solve.

The limitation of data has been identified as another problem with hypothetico-deductivism. The process of testing of a theory depends very much on the quality of the theory as well as that of the instruments that are used in the test to make observations. It is very difficult to conclude that our observations have falsified a hypothesis, because this may be due to a shortcoming in the experiment or the research design. It may be the result of a hypothesis which is not appropriate or based on some misconceptions; and/or faulty observations. This problem becomes even more crucial in the social sciences (see section on Hypothetico-deductivism in the social sciences).

Philosophers have argued that deduction is a logic of criticism, but not a logic of discovery or justification. The logic of discovery is to find new ideas or hypotheses, while the logic of justification is to support existing ideas or hypothesis. Deduction, although suitable for criticism, is less suitable

for discovery and justification. Criticism only works if the reasoning is deductively valid; thus the reasoning is not *ampliative* because it allows us to say that if the conclusion is false, then some premises must be false as well. In a valid deduction the conclusion is contained in the premises, therefore it does not aid in the discovery of new hypotheses. Similarly, it does not help to justify a hypothesis. Therefore, according to the critics of deduction, the logic of discovery and justification can only be nondeductive.

Even though there has been some criticism of hypothetico-deductivism, it has been the main framework used in scientific explanation. For most of the 20th century, it was seen as the default mode, or the only way, of doing science. It has been and still is successfully used in many fields as the main approach to scientific explanation, particularly in fields that have well-defined and falsifiable hypothesis which can be tested in controlled laboratory experiments, for example high-energy physics. However, not all fields (e.g., geology and cosmology) have such well-defined hypotheses which can be tested with laboratory precision. Similarly, hypothetico-deductivism is used in the social sciences, as discussed in the following section, but it requires some adaptation to deal with the nature of theory, testing, and operationalization of variables in those disciplines. In addition, the recent development of information technology and its rapid growth in the ability to collect and process large volumes of data have challenged the dominance of hypothetico-deductivism in the pursuit of scientific explanation.

Hypothetico-Deductivism in the Social Sciences

Hypothetico-deductivism has been used in the social sciences by many. The French sociologist Émile Durkheim's (1858–1917) famous study on suicide followed a deductive methodology. Arthur Stinchcombe (1933–2018) suggested an adaptation of Popper's falsificationism for the social sciences. On his view, research starts with theoretical statements related to the phenomena being studied. From these theoretical statements, operational definitions and empirical statements are deduced. Based on these empirical statements,

appropriate tests are designed and conducted. Based on the results of these tests, the theory is refuted or retained.

Stinchcombe uses Durkheim's theory of egoistic suicide to illustrate this point. According to this theory, a higher degree of individualism causes a higher rate of suicide in a group. From this theoretical statement, an empirical statement is logically deduced. For example, from Durkheim's theory of egoistic suicide some of the empirical statements which are derived are: Protestants in France have a higher suicide rate than Catholics in France and Protestant regions in Germany have a higher suicide rate than Catholic regions in Germany. Based on the empirical statements, a researcher can make observations according to the resources that are available to him them. Stinchcombe noted that theories often include empirical statements that are impossible to be observed given the resources which are available to the researcher. The empirical statements will also guide the operationalization of the variables in the study. Observations are used to test the theory. If the observations are found to be false, then one can prove that the theory is false. If the observations are corroborative, then the theory becomes credible. Sometimes, multiple tests are used to test a theory. If observations are confirmed in multiple tests, it helps the theory to become more credible than does a single test.

John A. Clark discusses a reconceptualization of hypothetico-deductivism for educational research. He discusses the logical relationship between three levels of explanation: observation sentences, observation categoricals, and theory formulations. A theory formulation is a sentence that comprises axioms of a theory. Observation categoricals are the empirical content of a theory formulation and are directly connected to observation sentences. Therefore, observation categoricals provide a logical connection between theory and observation. In discussing the use of hypothetico-deductivism in educational research, Clark argues that refuting a theory (in areas such as education) is not as simple as in classical syllogism. For example, given a theory, T predicts and observation, O , if the predicted O does not happen, then T will be rejected. However, if a conjunction of theories (T^1 – T^4) implies an observation O , when O is not

observed, which part of the conjunction needs to be revised or rejected cannot be determined by logic but by pragmatic judgment. Such pragmatic judgment is important when using hypothetico-deductivism in education.

Philip S. Gorski has criticized the use of hypothetico-deductivism in the social sciences, particularly in sociology. In hypothetico-deductivism, a single counterexample is enough to reject a theory, but, this hardly happens in the social sciences (or any sciences). No major theory in the social sciences has been rejected in this manner. One problem with hypothetico-deductivism in the social sciences is the notion of universal law which forms the major premise. Gorski argues that in sociology there has not been a universal law which can serve as the premise for deductive argument. He maintains that even the reformulation of hypothetico-deductivism in sociology will result in failure. Recognizing this problem, several social scientists have suggested adapted versions of hypothetico-deductivism. In social sciences, theory and observations are related and they influence one another. Theories can function as the lenses that are used to see the social world. The use of different theories will influence different interpretations of data. Since all observations are necessarily interpretations, we are not observing reality directly and theories cannot be refuted conclusively. Furthermore, some social scientists have argued that social theories are only capable of specifying the probability of certain outcomes. Hence, the theories are only falsified probabilistically.

Premalatha Karupiah

See also Deduction; Epistemology; Induction

Further Readings

- Blaikie, N. (2007). *Approaches to social enquiry* (2nd ed.). Polity Press.
- Blaikie, N. (2010). *Designing social research* (2nd ed.). Polity Press.
- Clark, J. A. (2000). Hypothetico-deduction and educational research. *Educational Research*, 42, 183–191. <https://doi.org/10.1080/001318800363827>
- Gorski, P. S. (2004). The poverty of deductivism: A constructive realist model of sociological explanation. *Sociological Methodology*, 34(1), 1–33. <https://doi.org/10.1111/j.0081-1750.2004.00144.x>

- Harrod, R. (1974). *Foundations of inductive logic*. Macmillan.
- Kvasz, L., & Zelenák, E. (2009). A problem for Popper's fallibilism. In Z. Parusniková & S. R. Cohen (Eds.), *Rethinking Popper*. Springer, 71–80.
- Mahootian, F., & Eastman, T. E. (2012) Complementary frameworks of scientific inquiry: Hypothetico-deductive, hypothetico-inductive, and observational-inductive. *World Futures*, 65, 61–75. <https://doi.org/10.1080/02604020701845624>
- Musgrave, A. (2004). Popper and hypothetico-deductivism. In D. M. Gabbay, S. Hartmann, & J. Woods (Eds.), *Handbook of the history of logic* (Vol. 10). Elsevier, 205–234.
- O'Hear, A. (1989). *An introduction to the philosophy of science*. Clarendon.
- Popper, K. (1959). *The logic of scientific discovery*. Hutchinson and Company.
- Stinchcombe, A. L. (1968). *Constructing social theories*. Harcourt, Brace and World.

INDUCTION

Induction is one of the two approaches of human reasoning or theory development identified by the philosophers of ancient Greece. (The other one is deduction, which draws out exceptionless consequences of given premises and is sometimes said to be an argument from the general to the particular.) Although some philosophers have proposed other ways of reasoning, such as abduction, the most popular textbooks of logic today continue to emphasize the two-approach theory of reasoning. This entry examines the definition of induction, its historical development, the problems associated with its use as method, its modern and contemporary forms, and the various models of induction used in scientific inquiry.

Induction Defined

The traditional definition of induction was proposed by Greek philosopher Aristotle (385–323 B.C.E.). His book *The Organon* (*The Instruments*) depicts induction as “an argument from the particular to the universal.” The prominent English philosopher John Stuart Mill (1806–1873) followed Aristotle’s thinking and defined induction, in his book *System of Logic*, as “the operation of the mind, by which we infer that what we know to be true in a particular case or cases will be true in all cases which resemble the former in certain assignable respects.” That line of definition remains in use by the influential *Oxford English Dictionary*,

which defines induction as “the process of inferring a general law or principle from the observations of particular instances (opposed to deduction).”

The logician Rudolf Carnap (1891–1970), however, classified all nondeductive reasoning as induction. This Carnapian approach holds that the only characteristic of induction is that its conclusions are not logically guaranteed by its premises. The premises only provide evidences to increase the probability of the truth of the conclusions. Widely used textbooks are normally seen as reflecting the broadest consensus among researchers in a field. Two such logic textbooks, Irving Copi and colleagues’ *Introduction to Logic* (in its 14th edition at the time of this writing) and Patrick Hurley’s *A Concise Introduction to Logic* (currently in its 13th edition), both use the broader Carnapian definition of induction and include analogy as a type of induction.

Proponents of induction claim that the idea of induction is to learn from experiences: people notice patterns and repeated phenomena in their lives and make projections about those patterns and phenomena in other places or in the future. Inductive scientists observe and measure individual objectives and phenomena, and then generalize those observations and measurements into educated guesses—in other words, hypotheses. Those hypotheses that are substantiated through sufficient testing and confirmation are called *theories* and are used to guide human understandings and behaviors. Australian geologist Karl Wolf, in a letter to the editor of the *International Journal of General Systems*, *Hypothesis versus Theory*:

Part of a Hierarchy of Truth (published in issue 5 of the journal in 2007), described the process as a journey along the sequence or hierarchy of truth, from problem, idea, concept/principle, hypothesis, to theory and eventually natural law, each latter step containing more reliability, accuracy, validity, and predictability than the former step. Other theorists or philosophers who place more emphasis on deductive thinking, such as Robert Dubin, author of the popular book *Theory Building*, define hypotheses as empirical indicators or predictions of theories, which can be observed and empirically tested. Hurley called the latter empirical hypotheses and the former theoretical hypotheses.

The Early Development of Induction

Although Aristotle probably provided the earliest definition of induction, he did not invest as much time and energy discussing induction as he did in discussing deduction and his famous syllogisms. Inspired by the development of Greek medical science, the philosopher Epicurus (341–270 B.C.E.) and his followers adhered to induction and argued that reasoning based on similarities among phenomena, that is, induction, was the only instrument for determining the laws of nature. In inductive reasoning, they believed, exhaustive listing of the members of a class was unnecessary. It should suffice if both similar and different facts were examined.

The statesman and philosopher Francis Bacon (1561–1626) has been widely considered as the most prominent advocate of induction. In his book *Novum Organum Scientiarum* (translated as the *New Instruments of Science*), Bacon criticized the deductive method, which had dominated the medieval era (represented by the Aristotelian syllogisms), as useless in scientific exploration. He argued that the syllogisms were not the fundamental principle of science but a tool to force people to agree with a conclusion. No new findings were obtained, as all information in the conclusion was contained in the major premise. Scientists' only hope, he said, was in induction. He further proposed three methods to identify causal relationships in phenomena, such as a listing table of presences and absences. Bacon is deemed the founder of the theory of experimental methods.

Charles Darwin (1809–1882), who established the theory of biological evolution, stated that his own work was based on the use of “the true Baconian method.”

John Stuart Mill (1806–1873) inherited Bacon's criticism of deduction and the idea of listing tables of presences and absences. He further developed his own five methods of scientific discovery to replace the unreliable traditional induction methods such as simple listing or incomplete listing. The five methods, designed to identify causal relationships, are described in his book *System of Logic*. The first, *method of agreement*, is used to find out the only factor that appears in all instances that the result appears (necessary cause). If in two or more instances the phenomenon under investigation had only one circumstance in common, wrote Mill, the circumstance in common was the cause of the phenomenon. When two or more circumstances are present in all instances when the phenomenon appears, however, the method of agreement cannot identify the cause. The second method, *method of difference*, may then be needed to find out the only circumstance that is absent in all instances that the result disappears (sufficient cause). The third method, the *joint method* of agreement and difference, is used in the following situation: in an instance the phenomenon under investigation appears and in another instance the phenomenon disappears, and both instances have all other circumstances in common except one, which is absent in the instance when the phenomenon disappears. The exceptional circumstance is therefore the (necessary and sufficient) cause of the phenomenon. Mill treated the joint method as a new and separate one, but later logicians have held that it is actually just a combination of the method of agreement and the method of difference. The *method of residues*, the fourth, tells researchers to find out all circumstances known not to be the cause of the result. The remaining circumstance then must be the cause. The method relies heavily on the accuracy of the existing literature that helps identify the circumstances known not to be the cause of the phenomenon under investigation. The fifth method, that of *concomitant variation*, is to find, from among a group of circumstances, the one that changes together with the result. The circumstance concomitantly varying with the result

is therefore the cause of the result. The method of concomitant variation needs to identify and get rid of the influences from the confounding circumstances that vary coincidentally with the cause.

Mill's five methods were later improved by English statistician Ronald A. Fisher (1890–1962). Fisher added statistical tests to them and turned them into the modern scientific method of experiment, which is still a powerful instrument in both natural and social scientific research. The key assumption of experiment and induction, said Mill, is the existence of a universal law behind the phenomena, such as the law of causation.

The Problems of Induction

In their debate with the Epicurean philosophers, the Stoic philosophers, who flourished in the Hellenistic period, developed their criticism of induction. They argued that induction was not a valid reasoning in form, which allowed human beings to inference from phenomena in their experiences to phenomena outside of their experiences. They pointed out that induction implied an assumption: hidden phenomena are similar to phenomena that were apparent or experienced. That assumption, serving as the major premise, actually turned induction into deduction. The Stoic philosophers further denounced induction as impossible to be conducted, because partial listing of the members of a class was insufficient and complete listing of those members was infeasible.

British philosopher David Hume (1711–1776) was the thinker who conspicuously drew attention to the problems of induction, which he stated in two forms: the logic problem and the psychological problem. He described the logic problem of induction with an example: because the sun had risen every day so far, one could not be absolutely sure that it would rise again the next day. People could not reason anything from instances repeated in their experiences to instances outside of their experiences, no matter how great the number of times the instances occurred in their experiences. Those instances in experiences, Hume believed, could not even help to make an inference on the probability of the other instances that people had not experienced. In the statement of the psychological problem, Hume explained the cause of the logic problem in the human mind.

The human being, he said, has a critical internal mechanism to generate associations on the basis of repetitions and maintain expectations of the instances that they have not experienced. Human beings cannot survive without such a mechanism, but the expectation produced through that psychological mechanism is not valid. Hume's criticism of induction, British philosopher Bertrand Russell pointed out, casts huge doubt on the description of the empirical procedures of scientific research. The question of induction actually led Hume himself to become a skeptical philosopher, doubting whether or not human beings truly have the capability to know the world.

Russell echoed Hume's questioning of induction in his *The Problems of Philosophy*, but it was Karl Popper (1902–1994), who made the questioning conspicuous again. Popper's criticism of induction was based on the Stoic point of view and Hume's logical and psychological problems of induction. The traditional question of induction, Popper said, is that no justification can be established for the belief that the future will be like the past. Without such justification, inductive inference is impossible. Also, to make an inductive reasoning valid, people must first establish an axiom of induction, stating something such as "Characteristics observed in many individuals in a population can be generalized to that population." Such a principle is hard to establish. Even if it is established, it then will be used to justify all inductive reasonings and therefore turns those inductive reasonings into deductive reasonings. Popper wrote that Kant's viewing induction as containing a priori validity was clever but nevertheless wrong to solve the problem of induction. Induction essentially states that the future is quite possibly the same as the past. Since it assumes too much about the future, it is actually not a reasoning method. The commonsense theory of knowledge, which assumes that everything enters our knowledge system through the human senses, was therefore invalidated for Popper. His solution was to totally reject induction and build a theory of scientific inquiry based only on deduction, that is, the method of falsification.

Popper's falsification method holds that universal statements cannot be generated from or proved by particular statements but can be rebutted by falsifying a particular statement deduced from it.

Scientists boldly propose hypotheses on issues of interest and develop theories. Good theories have more content and stronger explanatory power, while poor theories both contain and explain less. Scientists then test those good theories in an effort to falsify them. If a theory goes through a series of tests and cannot be falsified, it then enters the knowledge system. Induction therefore has no place in the theoretical construction method conceived by Popper. He went so far as to state that there is no such thing as induction.

Modern Forms of Induction

Despite Popper's total rejection of induction, modern logicians still uphold induction as a valuable tool of reasoning and scientific inquiry and continue to develop it. At the same time, current popular textbooks of logic adopt Rudolf Carnap's idea and expanded induction by putting all uncertain reasoning into this category. In Patrick Hurley's *A Concise Introduction to Logic*, for instance, induction includes not only the traditional inductive generalization from particulars to the general but also analogy, predictions about the future, arguments from authority, arguments based on signs (anticipating a sharp curve when seeing a sharp curve sign on the interstate), and so on. Hurley defines an inductive argument as one whose premises support the conclusions in a way that it is improbable (as opposed to impossible for deduction) for the premises to be true and the conclusions to be false. The common characteristic for all inductive reasonings, said Hurley, is that their conclusions all go beyond the contents of the premises.

Analogy is the inference from one particular to another particular that shares a reasonable amount of similarities with the former. The standard format of an analogy is as follows: A has attributes of a, b, c, and d; B has attributes a, b, and c; therefore, B probably also has attribute d. Since in most situations A and B should be in a category whose members all share those similarities, Hurley holds that analogy is closely related to traditional induction. The inference first unconsciously goes from A to that category along an inductive line, and then goes from the category to B, along a deductive line. The number of the similarities shared between A and B and the relevance of those

similarities determine the strength of an analogical reasoning. Legal research in common law systems, such as the law system in the United States or in the United Kingdom is, according to Hurley, essentially analogical. Legal arguments cite precedent cases that share most similarities with the case in question and request similar legal decisions. Moral reasoning, a process similar to legal research, is mainly analogical as well. Irving M. Copi and colleagues maintain that consumer decisions are also very much made through analogies. Consumers like to buy new products from manufacturers whose products they have used and felt satisfied. With analogies they reason that the new products should quite possibly also be satisfying. Every analogical inference, write Copi and colleagues, bases itself on one or more similarities of two things that serve as the subjects of the premises and the conclusion.

In their textbooks, both Copi and colleagues and Hurley label Mill's five methods as the second type of induction, which evidences that the five methods are not outdated in human reasoning since the derivation of the experimental method from them.

Probability and statistics are now well known as a *central topic of induction*. The probability theory states that when all events (the population) are independently and identically distributed (IID), or when every element in the population has exactly the same probability to be taken in the sample (probability sampling), the probability of one event (an element) in the population or sample happening can be calculated. Statistics are measurements of attributes in a sample of events taken from the population. Researchers then estimate the parameters in the population based on those statistics. The process of inferring from the sample (a group of particulars) to the population (the general) exemplifies the classical form of induction. Those statistical inferences contain errors, and the errors—called *sampling error in statistics*—are determined mainly by the sizes of the IID samples. Those inferences are also conducted at different levels of confidence. Higher levels of confidence are normally associated with more ambiguous predictions. In the social sciences, researchers usually choose 95% as the confidence level.

Statistical inference can be nonparametrical or parametrical. Nonparametrical statistical inference makes no other assumptions of the data

distribution beyond the IID assumption. Non-parametrical inferences are less sensitive to the information in the data, compared with their parametrical counterparts, and can only be used to infer simple relationships of a few variables. Parametric statistical inferences assume, in addition to the basic assumption of IID, that the data are normally distributed (generally containing more than 30 elements in a unit of analysis, and the data contain measurements at the interval or ratio levels). Parametric statistical tests can identify complicated relationships and are more sensitive to capture detailed information contained in the data. In modern social scientific research, advanced parametric statistical models such as structural equation models, hierarchical linear models, and multivariate time series models are often used to infer complex relationships from the samples to the populations. The truth of statistical inference is, of course, probable, at different levels of confidence, rather than absolutely guaranteed.

Although not yet included in many textbooks, the new concept of *big data* research is another form of induction. With the advancement of computing technology, an enormous amount of digital data can now be created and archived. The emerging research field of data science (also called *machine learning* in computer science, *data mining* in business or marketing research, and *big data* by many other people) analyzes those data, identifies patterns, and makes predictions. Big data analyses have comparatively few theoretical frameworks or hypotheses as guidelines for the analyses, and so the analyses are largely random-fishing and inductive in nature. The data, normally with their gigantic volume, are drawn from large, although perhaps not representative, samples or even from the populations. Big data analyses have been applied both academically and practically. In academic research, the combination of traditional social science research and data science analyses produces the area of computational social science research. In practical research, one of the major services of the leading database companies, such as Axiom, Epsilon, and Experian, is to provide big data analyses to help businesses better understand their customers' lifestyles, purchasing patterns, and needs in the near future. It is now common that new car purchasers receive car insurance mailers, or new house owners receive

furniture catalogues immediately after their purchases. Allegedly, a company identified a pregnant teenager by analyzing the commodities she had recently purchased and sent her mailers on baby products before her parents had even noticed the pregnancy.

The Inductive Models of Scientific Inquiry

Rudolf Carnap (1891–1970) is a major representative of the approach of induction. In Carnap's writing, natural sciences are described as *inductive sciences*. The process of scientific inquiry is a process in which scientists make precise observations (this step is viewed as the foundation of science), conduct experiments, and then generalize to build theoretical hypotheses and theories. In inductive inference, he said, scientists investigate a whole population based on one or more samples picked out of the population, possibly referring the statistical approach that is widely used in scientific studies today.

Inductive scientists always start from direct observations of particular facts, with the belief that only facts are available to observe. Patterns of the facts emerge in the observations and lead to conceptualization, definition, and classification (categorization) of those facts. The products of those conceptualizations—concepts—serve as bricks in the building of knowledge. Theoretical hypotheses connect the concepts and, after substantial testing, enter into the knowledge system as theories or laws, which provide guidance to understand other facts with similar patterns. An example of this inductive theory-building process is George Homans's sociological theory-building model, which contains three steps: First, sociologists observe how people behave in different groups; second, such behaviors are conceptualized and generalized; third, other groups are included in the study to confirm or improve the concepts and generalizations. Scientists in many different fields have embraced the inductive tradition, despite the major criticism that induction attempts to generate knowledge about the group, or the future, based on truth gained from particular elements of the group, or on the past. Some scholars believe that behaviorism, a school of psychology, also tends to use inductive generalizations based on substantial analyses of behaviors of the

particular subjects. Marketing researcher Gordon Foxall even claims that only inductive thinking is useful to study human behaviors.

Some social scientists adhering to inductive thinking embrace Carl Hempel's inductive-statistical (I-S) model. The I-S model is a derivation and modern version of Hempel's deductive-nomological (D-N) model. The D-N model consists of a law-like statement, an antecedent condition statement, and a conclusion deduced from the two premises. The I-S model simply changes the law-like statement into a statistical lawlike statement. Although deductive in its form, it is considered inductive by logicians because the truths of the premises, the statistical lawlike statement, and the antecedent do not guarantee the truth of the conclusion. The statistical lawlike statement, however, should state a relationship that has high probability to happen. An example of Hempel's I-S model could be:

Statistical law: If the volume of media coverage on an issue increases, people will very likely think of that issue as more important.

Antecedent: The volume of media coverage on education has increased.

Conclusion: People will very likely think of education as more important.

Wesley Salmon's statistical relevance (S-R) model is derived from Hempel's I-S model but has a more generalized form. Instead of claiming that the statistical lawlike premise should state a highly probable relationship, Salmon believes that the statistical lawlike statement can be any set of statements as long as they can provide a basis for inference of the probability of the conclusion, no matter how small the probability.

The invention of statistics, stated Ronald Fisher, had made induction as strict as, if not more than, deduction. The traditional thinking was that deduction was more accurate than induction because deduction only revealed the implications contained in the axioms, but induction generated new knowledge based on observations and deduction did not. Fisher, however, pointed out three factors that statistics made induction stricter than it had been thought previously. First, statistics led to more exact calculation of the observed data; second, the advancement of the experimental

method could generate better data for observation and calculation; and, third, statistics provided more sophisticated understanding from samples to populations or from the particulars to the general. Fisher also believed that it was induction, not deduction, that needed hypotheses as guidance to interpret the data.

Marketing researcher John O'Shaughnessy has described how induction is widely applied in scientific research in six steps: First, the scientists generate perceptual experiences about the topic under study by using observation. Second, by using definition and classification, the scientists sort out in order the facts that they have collected through observation. Third, inductive generalization is performed: The scientists make inferences (hypotheses) to a broader category based on the facts observed and analyzed. Then in the fourth step, the scientists test the generalized hypotheses by using observation again. In the fifth step, hypotheses that have withstood such testing are developed into theory, or law (the latter is more certain than the former), which in the next step become basis for further scientific explanations.

Copi and colleagues describe the process of scientific investigation in seven steps, mainly inductive but with deductive components. In the first step, the scientists identify the problem. In the second step, they develop preliminary hypotheses based on experiences and existing knowledge. In the third step, they collect additional facts, under the guidance of the preliminary hypotheses, to adjust and refine those preliminary hypotheses. In the fourth step, scientists formulate the formal explanatory hypotheses, which is ready to be tested. In the fifth step, scientists deduce from the explanatory hypotheses some testable consequences. In the sixth step, they test and observe the consequences, often in controlled laboratories. The explanatory hypotheses are supported if the consequences under test are consistent with predictions from those hypotheses. If the consequences under test are not consistent with the predictions, the explanatory hypotheses need to be revised and tested again. In the seventh step, explanatory hypotheses supported with intensive tests are treated as theories and applied in people's daily (or scientific) lives. In this process, Steps 2 to 4 are mainly inductive reasoning,

drawing general explanatory principles from particular facts observed by the scientists. In the fifth step, deduction is used to draw testable consequences from the explanatory hypotheses. In the sixth step, that of testing, however, induction (in the broader definition) is again used: the conclusion that the explanatory hypotheses are correct given that the predicted consequences are observed is probable, not absolute. This step's reasoning has been deemed logical fallacious by some classical logicians.

The role of induction in scientific research is also evidenced in the recent popularity of meta-analysis. Meta-analysis, by analyzing different sets of data collected in independent studies, aims to test a hypothesis or theory with more elements of data. Meta-analysis is inherently inductive and is used more and more by scientists as a strong tool in theory building or testing.

Qingjiang (Q. J.) Yao

See also Abduction; Big Data; Conceptual Analysis; Deduction; Falsifiability; Logic, Inductive

Further Readings

- Copi, I. M., Cohen, C., & McMahon, K. (2016). *Introduction to logic* (14th ed.). Taylor & Francis.
- Fisher, R. (1955). Statistical methods and scientific induction. *Journal of the Royal Statistical Society, Series B (Methodological)*, 17(1), 69–78.
- Hurley, P. (2017). *A concise introduction to logic* (13th ed.). Wadsworth.
- O'Shaughnessy, J. (1992). *Explaining buyer behavior: Central concepts and philosophy of science issues*. Oxford University Press.
- Popper, K. (1982). *Unended quest: An intellectual autobiography*. Open Court.

INFECTIOUS DISEASE STUDIES

Infectious diseases are disorders that are caused by microbial organisms that include bacteria, viruses, fungi, and parasites. These organisms are passed either directly or indirectly from person to person or animal to human. In fact, all eukaryotic forms of life (i.e., plants, animals, and humans) and even some bacteria (bacteriophages) are vulnerable to

invasion of an infectious organism that can lead to disease.

Infectious diseases have plagued humans since recorded history, and likely before. There are accounts of plagues (infectious diseases that infect large groups of people) during biblical times, and the Middle Ages (Black Death), the Spanish flu of 1918, and now COVID-19 in the 2020s. These plagues, also called *epidemics* (regional) and *pandemics* (global), have had major social and economic impacts on the affected populations, and with advances in modern travel technology, infections can more easily be transported globally, affecting millions in a very short period.

Statistics compiled by the U.S. Centers for Disease Control and Prevention (CDC) show that, over the last 35 years, mortality rates from infectious diseases in the United States have declined to approximately 19%. In 1980, there were approximately 41.95 deaths per 100,000 caused by infectious diseases, compared with 34.10 deaths per 100,000 in 2014. Eighty percent of infectious disease deaths were attributed to lower respiratory infections; however, there has also been a marked increase in diarrheal diseases, such as that of *Clostridium difficile* infections. HIV/AIDS is the third leading cause for infectious disease mortality, with 2.4 deaths per 100,000 individuals. There has been a shift of mortality from HIV/AIDS to the lower socioeconomic populations in the southeastern United States.

Even with significant advances in infectious disease research and therapies, the control and eradication of these diseases remain a daunting challenge. The World Health Organization has warned that infectious diseases are spreading more rapidly than ever before, and emerging infectious diseases are being discovered on a constant basis. Infectious diseases are difficult to combat for a myriad of reasons, including the emergence of new infectious diseases. Existing infectious diseases are spreading and distributing around the globe due to human activity or climate change or are being exploited for use in bioterrorism, antimicrobial resistance, and weakening public health infrastructures in countries with economic hardships. There are several major areas of research in infectious diseases, including vaccine research, hospital epidemiology and infection control, antibiotic resistance and stewardship, environmental infectious

diseases, and bioterrorism. Each of these areas of research converge around public health and safety.

Vaccine Research

The history of vaccines to be used in immunization of an individual or population most likely dates as far back as 1000 C.E. in China. The Chinese used material from a smallpox vesicle of an infected person to inoculate a healthy person. There are also accounts of inoculation with the virus in Africa and Turkey, which then spread to Europe and the Americas. However, in 1796, the English physician Edward Jenner tested cowpox to induce immunity to smallpox, which resulted in a milder form of the disease. Eventually, this practice became widespread and resulted in the eradication of smallpox. After this innovation, the French chemist and microbiologist Louis Pasteur developed the first rabies vaccine in 1885. This occurred around the time bacteriology was becoming defined as a discipline and resulted in many new developments, including vaccines against tetanus, anthrax, diphtheria, cholera, plague, tuberculosis, and typhoid.

By the middle of the 20th century, many new vaccines had been developed, and new methods for growing viruses allowed for discoveries and technologies, including the development of a polio vaccine. Following this, vaccines for measles, mumps, and rubella made great impacts on reducing the burden of these diseases.

Vaccines comprise some of the most cost-efficient preventive modalities in health care today. For every child that is immunized with vaccines on the current U.S. immunization schedule (DTap, Hib, MMR, Td, varicella, and Hep B), the effect is to save 33,000 lives and prevent more than 14 million cases of disease; furthermore, this entails a saving of about \$10 billion in direct care costs, as well as about \$35 billion in indirect costs. However, even with a well-planned vaccine schedule, roughly 42,000 adults and 300 children perish from vaccine-preventable diseases each year.

Today, there are many new techniques advancing vaccine research, with recombinant DNA technology and newer delivery methods. New disease targets are directing vaccine scientists to new paths to treat disease. One such example of a new delivery method is with the Severe Acute Respiratory

Syndrome or SARS-CoV-2 vaccine or mRNA vaccine delivery. Traditional vaccines use a microbial protein or an inactivated form of protein that induces an immune response in a host, whereas the mRNA is the genetic sequence of the SARS-CoV-2 virus spike protein that is reproduced by cells in the host. Both produce antigens that the body can use to stimulate the immune system; however, mRNA is faster to produce than the traditional vaccine. mRNA is noninfectious and is not a concern for DNA integration because it can't enter the nucleus where DNA resides. It is more cost-efficient than traditional vaccines that require expensive chemicals and cell cultures to produce. In a pandemic situation, such as that with SARS-CoV-2, the attainment of herd immunity, where enough individuals are vaccinated to protect those who are not, is the goal. Though great strides have been made to achieve herd immunity, there is still much work to be done to reach the number of vaccinated required to ensure adequate protection of the population.

Infection Prevention and Control

Epidemiology is the study of factors that determine disease occurrence in human and animal populations. Statistical methods are used to analyze epidemiological data (usually exposure-to-disease data) to attempt to understand and characterize disease outbreaks in different settings, including hospitals, cities, states, and nations. Clinicians, microbiologists, and other health care professionals (i.e., pharmacists and biomedical scientists) involved in preventative and public health care employ epidemiologic methods to gain understanding and manage various infectious diseases. Infection control in hospitals and other institutions such as prisons employ epidemiological principles to understand and control epidemics occurring at that level.

Infection prevention and control has its origins before the induction of germ theory, when in 1846, Ignaz Semmelweis, a physician from Hungary, showed that infections were caused by person-to-person spread in a labor and delivery unit in a major hospital in Vienna, Austria. To battle the infectious material being spread by clinical staff, a policy of handwashing was instated, using a chlorinated lime solution prior to contact with

patients. Additionally, John Snow, a physician based in London, intervened in a cholera epidemic in 1854 by removing a pump handle, preventing the intake of water from the pump (point source). Today, infection prevention and control departments operate in nursing homes and hospitals to derive policies, based on the basic principles of hand hygiene, appropriate handling of infectious materials, sterilization, and infectious disease surveillance to detect the beginning of a potential outbreak as quickly as possible.

Infections that are encountered in a hospital range from community-acquired infections such as methicillin-resistant *Staphylococcus aureus* (MRSA) and coagulase-negative *Staphylococcus aureus*, surgical site infections, urinary tract infections, and central line infections. Nosocomial infections (hospital-acquired infections) can occur, whereby clinical staff, in the absence of good hand hygiene practices, spread infections such as *Clostridium difficile* and multidrug-resistant organisms (MDROs) from patient to patient. Food-borne illnesses can also occur, as well as environmental causes (i.e., water towers breed *Legionella* spp.). In each of these infection types, the source and spread must be determined. Infection control practitioners use epidemiological principles to be able to determine the source of infections and employ the correct measures to get the spread under control.

An area in health care that frequently sees high levels of infection are with surgical patients. The National Health and Safety Network is a division of the CDC that has developed criteria that all infection prevention and control departments follow for the standardized identification and reporting of surgical site infections. In addition to reporting surgical site infections, routine monitoring is conducted within operating rooms and patient floors to ensure that proper hand hygiene and antiseptic technique is used when caring for these patients. Any findings of industry importance are analyzed and summarized in literature for the benefit of others within the field.

Environmental Infectious Diseases

The environment is often a significant contributing factor in the spread and occurrence of infectious diseases, both regionally and globally. Climate

change is emerging as a significant driver in the activity of infectious diseases, especially in the areas of vector-borne diseases, such as dengue, malaria, hantaviruses, and arboviral diseases. When examining the chain of infection, environment often has a key role in reservoir maintenance and route of transmission through air, water, and food.

During the 20th century, the world experienced large-scale anthropogenic changes in the environment. Intensive farming methods in conjunction with trade and travel are associated with the emergence of bovine spongiform encephalopathy (mad cow disease), foot-and-mouth disease, and viruses such as Nipah. Avian influenza is currently the greatest concern for global human health, because it can be transmitted from bird to human, as frequently happens in areas of Asia such as China or Indonesia. The sizes and proximities of flocks of poultry that are currently being kept for agricultural purposes would never occur in a natural setting. The proximities allow for the rapid spread of avian flu strains that can mutate and cause disease in humans more efficiently than in nature. Another example of infectious diseases spreading in the environment due to anthropogenic forcing is in Africa, where people hunt bushmeat that has had contact with many other types of animals carrying viruses. It is believed that bushmeat hunting accounts for the introduction of highly pathogenic viral diseases such as HIV/AIDS and Ebola.

Dam building and irrigation projects that oversee distribution of freshwater supply have introduced diseases such as schistosomiasis, Rift Valley fever, malaria, filariasis, onchocerciasis (river blindness), and Japanese encephalitis, to name a few.

It is important to understand these factors and how the environment plays a role in the emergence and spread of infectious diseases. Determining the causal links between climate change and disease emergence is difficult but requires a collaborative approach across states, provinces, and nations. It requires a variety of tools and expertise from many disciplines. Research involved in studying these outcomes will potentially result in better surveillance methods, detection, and collaborative response.

An example of research being conducted on infectious diseases and climate change is with hantaviruses. Hantaviruses are rodent-borne emerging

viruses. They cause two types of human disease: hemorrhagic fever with renal syndrome in Asia and Europe and hantavirus cardiopulmonary syndrome in the Americas.

Climate change, or global warming, has an impact on reservoir host behavior. Since most hantavirus infections occur in the warmer months, and rodent reservoirs are expected to migrate northward, the geographic distribution and frequency of infections is expected to increase in the next few decades. This phenomenon has already been observed in West-Central Europe, where there has been an increase in Puumala hantavirus outbreaks, in parallel with elevated average temperatures. Higher temperatures will lead to human activity that will increase exposures through aerosolized excreta.

Dams can be very useful to stabilize water availability and provide power but can also usher in disease outbreaks. In 1986, the Diama Dam was placed on the Senegal River to support agricultural expansion through water stability and prevention of saltwater intrusion. A few years after it was installed, a significant increase in human schistosomiasis, an infectious disease caused by a parasitic worm, was detected. The precise mechanism for the increase is not well understood but is believed to have been caused by the interaction of economic, epidemiological, and ecological factors. Dams have been associated with schistosomiasis for many years, and frequent problems occur around dams installed in Africa. Studies are being conducted to test the theory that introducing an abundance of *Machobrachium* species prawns will facilitate the natural reduction of schistosome snails that host the parasite, leading to lower incidences of disease in humans. This will help balance the need to use dams for power sources and control in areas endemic to the parasite.

Antibiotic Resistance and Stewardship

Antibiotic resistance is a very important problem in the study and control of infectious diseases. It is a significant contributor to mortality and morbidity, prolonged hospital stays, and increased health care costs. Simply put, overuse and inappropriate use of antibiotics results in the development of antibiotic resistance. The vast majority of

antibiotic use occurs in primary care settings and nursing homes, but it is a significant problem in hospitals, where MDROs are most likely to develop. The CDC has classified several bacteria as causing urgent, serious, and concerning threats. These resistant bacteria, such as MRSA, ESBL, and vancomycin-resistant enterococci, result in increased health care costs. Antibiotic stewardship is a means to direct and offer guidelines for the judicious use of antibiotics in hospitals to help reduce or eliminate the problem of antibiotic resistance.

Antibiotics first emerged with the discovery of penicillin by Sir Alexander Fleming in 1928. Penicillin was very effective in controlling infections among World War II soldiers. Yet, shortly after the war, penicillin resistance began to emerge. By 1962, the first case of MRSA was identified in the United Kingdom, followed by the United States in 1968. Now, resistance to nearly all antibiotics has been observed in the clinical setting. From the 1960s to 1980s, the pharmaceutical industry introduced new antibiotics, but subsequently little research and development has been implemented.

Bacteria acquire resistance to antibiotics by inheritance from relative bacteria or through non-relative bacteria by a circular DNA organelle called a *plasmid*. When acquired by a plasmid, a horizontal gene transfer causes antibiotic resistance to be passed between different species of bacteria. Resistance can also occur spontaneously through mutation. Antibiotics kill drug-sensitive bacteria that would compete for resources, resulting in the proliferation of resistant bacteria.

One of the main drivers of antibiotic resistance includes inappropriate use, or clinicians prescribe an antibiotic without investigating the species type or sensitivity of the bacteria to a particular antibiotic. Frequently, infections for which antibiotics are prescribed are actually viral and not bacterial in nature; such infections cannot be treated by antibiotics. Additionally, antibiotics that are not prescribed in a therapeutic dose but in a dose too low to be effective can promote the development of antibiotic resistance through supporting genetic alterations. These genetic alterations can lead to increased virulence and strain diversification.

Another way that antibiotic resistance has occurred is through extensive agricultural use. Globally, antibiotics are prescribed as a means of

growth enhancement in livestock. Approximately 80% of antibiotics sold in the United States are employed in livestock production to prevent infection and promote growth. It is said that treating livestock with antibiotics can lead to healthier animals and increase larger and higher quality yields. However, humans ingest the unmetabolized antibiotics when they consume animals that are treated with large quantities of the drug.

In addition to the negative effects on animal and human health, the overuse of antibiotics affects the environmental microbiome. Nearly 90% of antibiotics administered to livestock are excreted through waste and deposited in fertilizer, groundwater, and surface runoff. This inevitably increases spread of resistant organisms within the environment.

Antibiotic stewardship programs (ASPs) oversee measures within a hospital that promotes judicious use of antibiotics. They are developed at the hospital level; there is no standardized, universal set of rules for the use of antibiotics at this time. A growing number of research articles within the ASP arena indicate that antibiotic stewardship is a priority, but several research gaps still need to be addressed. Barriers to the development of ASPs need to be identified, as well as the correct actors implementing change; in addition, more work is needed on translating research into stewardship guidelines, defining outcomes, establishing an evidence-based approach for the selection of specific antibiotic regimens, and understanding the role of governmental policies on ASPs.

Research that explores ways to slow or eliminate antibiotic resistance is under way. Collaborative efforts, such as the Consortium on Resistance Against Carbapenems in *Klebsiella* and other Enterobacteriaceae (CRACKLE-1), headed by David van Duin, MD, PhD, at the University of North Carolina, seeks to understand the mechanism of resistance for carbapenem-resistant Enterobacteriaceae (CRE). Results from the CRACKLE-1 consortium showed that ST258, the primary strain of CRE *Klebsiella*, has subdivided into various strains that have distinct clinical outcomes. This data led to the creation of CRACKLE-2, which had over 60 hospitals in participation. CRACKLE-2 was a multicenter research study that collected clinical and epidemiological data on patients diagnosed with MDRO infections to

assess the desirability of outcome ranking measured at 30 days from index culture. It also characterized the MDRO species, strain type, and resistance mechanism to carbapenem. Research such as CRACKLE-1 and 2 will yield information that can help direct hospital Antibiotic Stewardship programs and infection control practitioners to more adequately address the proclivities of CRE antibiotic resistance and improve patient outcomes while decreasing hospital costs.

Bioterrorism

Although not a new concept, bioterrorism, or the use of a biological agent to incapacitate or kill another person or group of people, is an important topic in the area of infectious disease studies. Biological agents can be spread by air, water, or food. Once infected, a victim can spread the disease to another person. Biological warfare is insidious because it is hard to detect and usually does not cause illness for several hours or days after being released.

There are several incidences of bioterrorism where an infectious disease was used during wars, as early as 600 B.C.E. In that period, people catapulted cadavers, excrement, and animal carcasses riddled with germs into enemy camps to incapacitate the enemy and win conflicts. Pollution of wells occurred in several European wars, the American Civil War, and also in the 20th century. The Japanese military used plague as a biological weapon by inducing transmission of the disease through infected fleas to rats in a laboratory setting. They then released plague-infected rats in Chinese cities to initiate an epidemic. This backfired, though, and about 1,700 Japanese troops died after being exposed. German medical researchers infected prisoners with *Rickettsia prowazekii*, malaria, and hepatitis A virus at the time of World War I. In the United States, a biological warfare program was initiated in 1942 by the War Reserve Service. Today, the government entity responsible for bioweapons research and development is at Fort Detrick, Maryland (U.S. Army Medical Research Institute of Infectious Diseases [USAMRIID]).

In 1972, the United Nations General Assembly adopted the Convention on the Prohibition of the Development, Production and Stockpiling of

Bacteriological (Biological) and Toxin Weapons and on Their Destruction (BWC). Under this treaty, any development, production, or stockpiling of pathogens or toxins among member countries is prohibited. It required all parties involved to destroy any existing stockpiles, delivery systems, or equipment within nine months of the treaty being signed. In the United States, President Richard M. Nixon signed an executive order in 1969 and 1970 to terminate the offensive biological weapons program. Along with this order, the United States created a policy to ban the use of biological weapons, including toxins (e.g., botulism). The only caveat was to allow research efforts to continue for countermeasures to potential attacks, including vaccines and antisera. Between May 1971 and February 1973, the entire stockpile of biological weapons was destroyed, overseen by the U.S. Department of Agriculture, Nature Resources of Arkansas, Colorado, and Maryland. USAMRIID continues to conduct research for development of medical defense for the military in the event of an attack. The research conducted is considered open and is not classified.

The threat of a biological attack is always present, and the medical community, as well as public health organizations, must be diligent in surveillance, epidemiology, and countermeasures for mass hysteria if a biological weapon is used on the population. To understand the effects of biological weapons on a population, modeling is used. These models look at the worst-case scenarios of exposures for millions of human casualties or fatalities. Though simulations are not perfect solutions, they do help form hypotheses and raise questions on preparedness that could not be done without the models.

Conclusion

There are many areas in health care and the environment with which infectious diseases intersect, and ongoing research is necessary to continue to educate not only people in health care and science fields but also laypersons. A thorough understanding of the pathophysiology and mechanisms of infectious diseases is the foundation for all other endeavors, such as vaccine research, antimicrobial resistance, environmental health, infection

prevention and control, and bioterrorism research. Though all these areas are important, antimicrobial resistance and vaccine research are of the highest priority because quarantine and nonmedical control measures are effective, but without a means to treat infections, deaths would continue and potentially be uncontrollable.

Mandy M. McBroom

See also Experiment, Theory of; Health Care Science; Life Sciences, Historical; Medicine, 19th Century

Further Readings

- Brock, T. D. (1988). *Robert Koch: A life in medicine and bacteriology*. American Society for Microbiology.
- Bynum, W. F. (1994). *Science and the practice of medicine in the nineteenth century*. Cambridge University Press.
- Centers for Disease Control. (1999). Ten great public health achievements—United States 1900–1999. *MMWR*, 12, 241–243.
- Conrad, L. I., Neve, M., Nutton, V., Porter, R., & Wear, A. (1995). *The Western medical tradition 800 BC to AD 1800*. Cambridge University Press.
- Llah, S. T., Mir, S., Sharif, S., Kahn, S., & Mir, M. A. (2018). Hantavirus induced cardiopulmonary syndrome: A public health concern. *Journal of Medical Virology*, 6, 1003–1009.
- Opal, S. M. (2009). A brief history of microbiology and immunology. In A. W. Arstenstein (Ed.), *Vaccines: A biography* (pp. 31–56). Springer-Verlag.
- Soklow, S. H., Jones, I. J., Jocque, M., La, D., Cords, O., Knight, A., Lund, A., Wood, C. L., Lafferty, K. D., Hoover, C. M., & Collender, P. A. (2017). Nearly 400 million people are at higher risk of Schistosomiasis because dams block the migration of snail-eating river prawns. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 1722, 20160127.
- Torriani, F., & Taplitz, R. (2010). History of infection prevention and control. In J. Cohen, S. M. Opal, & W. G. Powderly (Eds.), *Infectious diseases* (3rd ed., pp. 76–85). Mosby.
- Van Dubin, D., Arias, C. A., Komarow, L., Chen, L., Hanson, B. M., Weston, G., Cober, E., Garner, O. B., Jacob, J. T., Satlin, M. J., & Fries, B. C. (2020). Molecular and clinical epidemiology of Carbapenem-resistant Enterobacterales in the USA (CRACKLE-2): A prospective cohort study. *The Lancet Infectious Diseases*, 19, 30755–30758.

INFERENCE

Inferences drive scientific inquiry. Inferences are instances of reasoning that take some given evidence as input and yield, as output, a judgment, or statement. Prominent outputs for scientific inference concern statistical associations, cause–effect relations, and the pursuit worthiness of research programs. Especially famous inferences include Euclid’s proof of the Pythagorean theorem from various postulates and diagrams, Archimedes’ (legendary) discovery of the law of buoyancy while observing his body displace water from a tub, Albert Einstein’s inference from a thought experiment about a man falling freely through space inside an elevator to the equivalence of gravitational mass and inertial mass, and Alpha-Go’s strategic inferences to defeat grandmaster Lee Sedol in four games of a five-game Go match.

Categorizing Inference

Attending to divisions among classes of inference facilitates understanding the nature of inference. Three prominent categorical frameworks divide inferences into deductive and inductive, demonstrative and ampliative, and formal and material.

Deductive Versus Inductive

Contemporary logicians typically divide the class of inferences into deductive inferences and inductive inferences. With deductive inferences, the inputs to reasoning, or premises, purport to guarantee the truth of the outputs or conclusions. With inductive inferences, the inputs only purport to make the outputs more likely than not. For example, despite the standard terminology, the proof by mathematical induction that $11^n - 6 = 5m$ (for every positive integer n and some integer m), from the base case $11^1 - 6 = 5(n=1, m=1)$, is a deductive inference. By contrast, in their 1977 article “Every Planar Map is Four-Colorable,” the mathematicians Kenneth Appel, Wolfgang Haken, and John Koch present their computer-assisted proof of the four-color theorem from the axioms of graph theory as inductive, by virtue of its appeal to (likely harmless) probabilistic assumptions.

Deductive inferences are good (valid) when the truth of their premises precludes the falsity of their conclusions, and bad (invalid) when they do not. For example, the standard proof for the inequality of 1 and 0 from Peano’s axioms for arithmetic is a valid deductive inference, because the truth of those axioms guarantees that $0 \neq 1$. Inductive inferences are good (cogent) when the truth of their premises makes likely the truth of their conclusions and bad (uncogent) when they do not. For example, Isaac Newton’s inference, in his *Principia Mathematica*, from Proposition 5 of Book 3 (that the moons of Jupiter gravitate toward Jupiter and the moons of Saturn gravitate toward Saturn) to Corollary 1 thereof (that there is gravity toward all other planets) is a cogent inductive inference, because it depends upon the (correct) assumption these other planets also have moons orbiting them.

Demonstrative Versus Ampliative

The distinction between deductive and inductive inference resembles another distinction, that between demonstrative and ampliative inference. Inferences are demonstrative when their inputs contain all of the content or information in their outputs (and perhaps more). They are ampliative, by contrast, when the content of their outputs amplifies, or exceeds, that of their inputs. Because the premises of demonstrative inferences necessitate their conclusions, all good deductive inferences are demonstrative. There is reason to suppose that some inductive inferences are also demonstrative. For example, so-called Holmesian inferences, which infer that some theory must be correct because all alternative theories are false, seems to be both inductive, by virtue of the implicit assumption that all possible alternatives are known, and also demonstrative. Many other inductive inferences, such as Newton’s inference from Proposition 5 of Book 3 to Corollary 1 thereof and Appel, Haken, and Koch’s inference from the axioms of graph theory to the four-color theorem, are ampliative.

Although the preceding distinctions, between deductive and inductive inference and between demonstrative and ampliative inference, are helpful for categorizing specific instances of reasoning,

they are less than helpful for understanding how scientific reasoning works in practice. Three observations are salient in this regard. The first is that the preceding distinctions ignore inferences that invoke inputs that lack truth-values, because they presume that the inputs for good inferences are always declarative statements. This presumption has the unfortunate consequence, for example, that Euclid's proof of the Pythagorean theorem is a bad inference, because Euclid invokes various (nondeclarative) commands and visual intuitions about particular geometrical diagrams when presenting his proof.

The second observation, derived from the same reason as the first, is that the preceding distinctions also ignore inferences that yield outputs that lack truth-values. Yet some scientific inferences yield, as outputs, recommendations to pursue one research hypothesis or strategy rather than another. For example, given two competing models that perform equally well with respect to available experimental data, scientists sometimes infer that one model is more plausible than the other by virtue of being more robust to alterations of its parameters, their plausibility judgment is meant to recommend the one model as worthier of further research efforts—and not to recommend that the one model is more likely to be true than its competitor.

The third and final observation is that adding as input a claim to the effect that *if the inputs are correct*, then the output is correct suffices for converting any inductive inference into a deductive one and any ampliative inference into a demonstrative one. Applying the distinctions to categorize various inferences thereby depends heavily upon decisions about whether to interpret those inferences as containing implicit claims that secure their validity.

Formal Versus Material

For the sake of better understanding how scientific reasoning works in practice, contemporary logicians distinguish between formal inferences and material inferences. An inference is *formal* when its goodness depends entirely upon its logical form. Such inferences conform to formal schemes. Paradigmatic examples include *modus ponens*, which has the form *if P, then Q; P;*

therefore Q, and statistical syllogism, which has the form *N% of individuals in reference class F also belong to attribute class G; individual x belongs to F; therefore, x also belongs to G*. An inference is *material*, by contrast, when its goodness depends at least in part upon its empirical content or background knowledge. Inductive generalizations, or inferring that some characteristic of a sample holds of a broader population, are often material inferences. Consider, for example, the inference from observations that some samples of a new organic salt melt near a certain temperature and standard pressure to the conclusion that all such salt melts near the same temperature. This inference seems to conform to a formal scheme with an implicit premise that whatever is true of a representative sample is true of the corresponding broader population. But the inference is material, because determining whether the observed samples of salt are representative requires an appeal to background knowledge about variability among organic compounds of the same kind.

Classifying some inferences as material inferences helps to explain several features of scientific theorizing in practice. It helps to explain why conceptual change sometimes affects the goodness of inferences by focusing attention on the empirical content of premises. Consider, for example, the predictive inference from *this pea plant has a certain gene* to *this plant's peas will be a certain color*. This is a cogent inference scheme in the context of classic Mendelian genetics, where the concept of *gene* refers to an allele that determines an organism's appearance. But it is an uncogent scheme in the context of contemporary developmental biology, where the concept of *gene* refers to particular DNA nucleotide sequences and does not determine an organism's appearance. Classifying some inferences as material inferences also helps to explain how scientific discovery can be rational without proceeding in accordance with general rules of inquiry. Consider, for example, Hans A. Krebs's discovery of the urea cycle that converts toxic ammonia to excretable urea. Krebs's discovery relies upon stoichiometric considerations that are particular to metabolic chemistry, such as criteria for specifying the sequence of metabolic reactions and for balancing chemical equations. Insofar as the inferences responsible for his discovery are material, it is background

knowledge about metabolic chemistry, rather than any formal schema, that licenses the cogency of Krebs's inferences.

Varieties of Inference

Inferential strategies in the sciences, engineering, and mathematics are many and varied. What follows is a brief review of some strategies that are especially prominent and important in contemporary scientific practice.

Statistical Inference

Statistical inferences appeal to principles of probability or statistics. The two most dominant traditions for conceptualizing statistical inference are *frequentist statistics* and *Bayesian statistics*. Both traditions agree that likelihoods—probabilities of obtaining some result or output given the presumed truth of some hypothesis—are relevant to making inferences as the factors responsible for those results or outcomes. They disagree about which likelihoods are relevant and about how relevant likelihoods function as premises for inference.

Frequentist Statistics

Frequentist statistics, also known as *classical statistics*, involves inferring from probability statements about real or hypothetical long-run frequencies pertaining to a data sample from a population to conclusions about whether one ought to accept or reject various hypothesis about the populations from which those frequencies derive. In the social and behavioral sciences, the dominant paradigm for frequentist statistical inference is *null hypothesis significance testing* (NHST). NHST begins by specifying a null hypothesis, H_0 , to the effect that some factor makes no difference to a phenomenon of interest or that some intervention has no effect upon a system of interest. Paradigmatic examples of null hypotheses include the hypothesis that a coin is fair (and thereby does not contribute to whether flipping the coin yields *heads* or *tails*), and the hypothesis that some medicine does not reduce the symptoms of some disease. NHST contrasts the null hypothesis with a contrary alternative hypothesis, H_A , according to which the factor or intervention of

interest has a statistically significant impact upon the outcome of interest.

Given a data sample obtained from observation or experimentation, NHST proceeds by selecting a test statistic for the data—such as mean data value or variance among data values—and calculating the value of this statistics from the sample data. NHST determines the evidential bearing of this statistic upon the null hypothesis by selecting a testing rule. Testing rules determine how likely a null hypothesis should make the value of the test statistic in order for the data to not count as evidence against the null hypothesis. Each rule relate a p value to an *alpha* level, where the p value is the probability, relative to the null hypothesis, of obtaining a value for the selected test statistic that is at least as extreme as the actually obtained value, and the alpha level is some number α —typically 0.05 or 0.01—such that data count as evidence against the null hypothesis if and only if the p value for that data is not greater than α . Calculating the p value for a given data sample then permits an inference either to the conclusion that the null hypothesis is likely correct (if $p > \alpha$) or to the conclusion that the alternative hypothesis is likely correct because the null hypothesis is likely incorrect (if $p \leq \alpha$).

Whether inferences that derive from NHST are reliable is a matter of ongoing debate. So, too, are views about details of the inferential procedure for NHST. NHST is an amalgam of competing inferential procedures. The reliance upon p values derives from a Ronald Fisher's method for determining whether associations among data are due to chance. The reliance upon alpha values when testing for significance, derive from Jerzy Newman and Egon Pearson's method for determining whether prudence counsels accepting or rejecting hypotheses. Standard criticisms typically fault NHST-based inference for being incoherent, lacking principles that determine whether a test statistic is appropriate, and correlating low p values with high degrees of support for alternative hypotheses.

Bayesian Statistics

Bayesian statistics involves inferring from probabilities pertaining to a data sample D and various hypotheses to a conclusion about the degree to

which the data supports one of those hypotheses. These inferences involve three substantive premises: a statement about the prior probability of some hypothesis of interest, $\Pr(H)$, which indicates how likely the hypothesis is in relation to existing evidence; a statement about the probability of the data in relation to the hypothesis of interest, $\Pr(D|H)$, which indicates how likely it would be for the data to obtain if the hypothesis were true; and a statement about the probability of the data itself, $\Pr(D)$, which is the straight sum of products between the prior probability of one of several exclusive and exhaustive hypotheses H_i and the probability of the data in relation to that hypothesis (such that $\Pr(D) = \sum_i \Pr(D|H_i) \times \Pr(H_i)$). These premises support a conclusion about the degree to which the data support the hypothesis of interest, $\Pr(H|D)$, by virtue of a theorem from probability calculus known as *Bayes' Rule*.

$$\Pr(H|D) = \frac{\Pr(H) \times \Pr(D|H)}{\Pr(D)}$$

Consider how Bayesian statistical inference compares to NHST-based inference in the case of determining whether a coin is fair. Let the null hypothesis be that the coin is fair: $H_0 := \Pr(\text{heads}) = 0.50$. Suppose background evidence excludes all but one alternative hypothesis, According to which the coin is slightly biased toward heads: $H_A := \Pr(\text{heads}) = 0.55$. (Perhaps the coin is known to originate from one of two machines.) Let the relevant experiment be flipping the coin 1,000 times. Let the data show that the coin lands on heads exactly 528 times. Then $\Pr(D|H_0) \approx 0.0056$ and $\Pr(D|H_A) \approx 0.0095$. Supposing that the null and alternative hypotheses are equally likely, $\Pr(H_0) = \Pr(H_A) = 0.5$. Since $\Pr(D) \approx 0.0076$, Bayes' Rule entails that $\Pr(H_0|D) \approx 0.3712$ and $\Pr(H_A|D) \approx 0.6288$. Bayesian statistical inference therefore concludes that the coin is more likely biased than fair. The NHST procedure, by contrast, relativizes inference to a testing rule. Let the rule be that H_0 is likely correct if and only if the p value for the experiment is not greater than 0.05. Since the p value equals $\Pr(D|H_0) = 0.0056$, the NHST procedure entails that there is no significant disagreement between the data and the null hypothesis, and so the coin is likely fair. (If the α value were selected to be 0.01 rather than 0.05,

or if the coin had landed on heads more than 528 times, the same procedure would entail that the coin is likely biased.)

Causal Inference

Statistical inferences are vehicles for theorizing about statistical associations among factors in some system or population of interest. However, they do not yield reliable conclusions about causal connections. Statistical inferences lack power to discriminate between one factor causing another and one factor being coincident effect of some other common cause. They also do not address issues about what would happen to a system or population if some putative cause were different. Several approaches to causal inference aim to overcome these limitations of (mere) statistical inference. The two most dominant approaches to causal inference are the potential outcomes approach (POA) and the structural modeling approach.

Potential Outcomes Approach

POA involves inferring from experimental data about a system of interest to conclusions about the causal effect of one factor relative to another. The experimental data typically derive from comparing the effects of a test intervention upon some factor of interest to the effects of an alternative control intervention, and the causal inference is a matter of calculating the causal effect of the test intervention relative to the control as a difference between the effects of the test and control, respectively, after multiple interventions. For example, if the system of interest is the human body and the effect of interest is alleviation of some bodily ailment, if a sample of people with headaches who take a specific medicine typically have headache relief within an hour while a sample of people with headaches who take no medicine typically have no relief within an hour, and if the samples of people are otherwise similar, POA might infer that taking the medicine has a causal effect relative to not taking any medicine. (Determining whether the inference is warranted typically involves using frequentist statistics to determine whether a causal effect is statistically significant.)

Whether data warrant an inference about causal efficacy depends, according to POA, upon a variety

of considerations. The calculated causal effect should be statistically significant. The composition of the test and control samples should be unbiased. The samples should not differ with respect to other factors that might be causally efficacious. They also should represent the population or system of interest. Frequentist statistical inference with respect to a null hypothesis about expected causal effect, random sampling from a population of interest, and randomization with respect to potentially confounding factors are standard tactics for satisfying these desiderata.

Structural Modeling Approach

The structural modeling approach to causal inference involves using causal hypotheses to build a statistical model, testing the hypotheses in the model, and evaluating test results to infer to a conclusion about whether the model is valid. Causal hypotheses are speculations about cause-effect relations among different factors within a population or system of interest. These speculations are based upon background knowledge and available correlational data, and combining them yields a model for a structure of cause-effect relations. Testing a model involves various statistical methods, such as parameter estimation and goodness of fit, to determine how well the model describes available data. Such tests aim to determine whether a model—and its associated causal hypotheses—is valid in one of four senses, namely, whether the model supports claims about covariation among various factors, whether it represents causal relations, whether alternative cause-effect constructs yield consistent results, and whether the model generalizes to other populations or systems.

Heuristics

When the need to make inferences is pressing, limitations of time, knowledge, and computational power favor fast and frugal inferential strategies over provably reliable ones. These strategies often take the form of heuristics. Because the reliability of these heuristics—their capacity to produce good inferences rather than bad ones—often depends upon the context in which they occur, the heuristics are often strategies for material inference and therefore relatively limited

in scope. Consider, for example, the task of programming a computer capable of inferring appropriate moves for games of chess. The rules of chess impose time limitations for completing each game of chess, and processor capacities limit the speed and number of calculations any computer is able to perform. Because of these limitations, modern chess programs typically rely upon heuristics that infer appropriate moves by evaluating only some available information about the chessboard, such as the relative value of remaining pieces, the number of attacking pieces in the board's center, and the number of pieces in position to attack the opponent. The success of these heuristics depends not only upon the astounding computational powers of modern computers but also upon the high number of constraints that chess imposes in comparison to, say, the constraints relevant to programming an autonomous vehicle to navigate dense urban traffic—such as a relatively small environment, a relatively limited number of moves all of which have well-defined consequences, and a single goal.

Even when the need to make inferences is not pressing, there are reasons to favor inferential strategies that are not directed toward correctness. Inferences at the frontiers of inquiry often are directed toward discovering appropriate next steps for research rather than producing conclusions that are likely to be true. Such inferences require ranking the potential for pursuing one line of research rather than another. Because research frontiers are typically awash with uncertainty, ranking competing courses of research often involves heuristic strategies for inferring that much of the information available about a research topic is likely irrelevant. For example, rather than read thousands of article abstracts, one might use the heuristic *only read articles that contain specific keywords* to infer that many available articles in one's research discipline are likely irrelevant to one's research project. Similarly, rather than attempt to justify whether a particular research procedure is worth developing, one might use the heuristic *develop any research procedure that produces surprising results*.

Because research funding is limited and research efforts devoted to one line of inquiry preclude devoting attention to others, determining whether a model or theory at a research frontier is worthy

of further development also often involves heuristic strategies of inference. Consider, for example, a brief history of the liquid drop model for atomic nuclei. In the late 1920s, the Russian physicist George Gamow proposed to treat atomic nuclei as drops of water whose constituent particles are held together by surface tension. Gamow used his model to explain the nuclear binding energies of different nuclei, but the model's predictions were only accurate for relatively light elements. Although Gamow abandoned his model, several European physicists continued efforts to revise and extend it. These efforts eventually led to the discovery of the semiempirical mass formula as well as explanations for artificially induced radioactivity and nuclear fission. But the physicists had no way to know, in advance, that their efforts would be fruitful, because the limited success of a simple model does not guarantee that revisions or extensions of the model also will be successful. Instead, their decision to develop Gamow's model likely relied on a heuristic inference from the fact that adapting an existing and well-understood modeling strategy to a new domain is more cost-effective than developing a novel or poorly understood strategy to the conclusion that Gamow's model was worth pursuing.

Nicholaos Jones

See also Abduction; Deduction; Hypothesis Testing; Induction; Inference to the Best Explanation; Logic, Formal and Informal; Reasoning; Statistical Inference, Bayesian; Statistical Inference, Frequentist

Further Readings

- Appel, K., Haken, W., & Koch, J. (1977). Every planar map is four-colorable. *Illinois Journal of Mathematics*, 21(3), 429–567.
- Barnett, V. (1999). *Comparative statistical inference* (3rd ed.). John Wiley & Sons.
- Brigandt, I. (2010). Scientific reasoning is material inference: Combining confirmation, discovery, and explanation. *International Studies in the Philosophy of Science*, 24(1), 31–43. <https://doi.org/10.1080/02698590903467101>
- Chow, S. L. (1996). *Statistical significance: Rationale, validity, utility*. Sage.
- Gigerenzer, G., Todd, P. M., & ABC Research Group. (1999). *Simple heuristics that make us smart*. Oxford University Press.
- Howson, C., & Urbach, P. (2006). *Scientific reasoning: The Bayesian approach*. Open Court Publishing.
- Hubbard, R. (2004). Alphabet soup: Blurring the distinctions between *p*'s and *α*'s in psychological research. *Theory & Psychology*, 14(3), 295–327.
- Imbens, G. W., & Rubin, D. (2015). *Causal inference for statistics, social, and biomedical sciences: An introduction*. Cambridge University Press.
- Jones, N. (2018). Inference to the more robust explanation. *British Journal for the Philosophy of Science*, 69(1), 75–102.
- Nickles, T. (2016). Fast and frugal heuristics at research frontiers. In E. Ippoliti, F. Sterpetti, & T. Nickles (Eds.), *Models and inferences in science* (pp. 31–54). Springer.
- Norton, J. (2003). A material theory of induction. *Philosophy of Science*, 70(4), 647–670.
- Nyrup, R. (2020). Of water drops and atomic nuclei: Analogies and pursuit worthiness in science. *British Journal for the Philosophy of Science*, 71(3), 881–903. <https://doi.org/10.1093/bjps/axy036>
- Mouchart, M., Russo, F., & Wunsch, G. (2010). Inferring causal relations by modelling structures. *Statistica*, 70(4), 411–432.
- Pearl, J. (2000). *Causality: Models, reasoning, and inference*. Cambridge University Press.

INFERENCE TO THE BEST EXPLANATION

Inference to the best explanation (IBE) is a type of scientific reasoning whereby the best explanation for a scientific phenomenon is accepted as the correct explanation. IBE stems from the belief that causes cannot be known with absolute certainty but are known only through highly probable theories about causality. IBE is distinct from deductive and inductive reasoning and is often considered a third type of reasoning associated with *abduction*, set forth by philosopher Charles Sanders Peirce. The terminology IBE was coined by Gilbert Harman to prevent misunderstanding as to what was being reasoned or inferred. In particular, IBE proposes that it is not the cause, as such, that is being inferred but rather the explanation for the effect. This type of reasoning happens in reverse order from basic inductive reasoning.

Terminology

Harman coined the term *inference to the best explanation* in 1965 to propose that inferences in scientific study are all extensions of one basic point, which is that humans infer reasons for effects. Harman was not the first to put forth this notion but drew on others who have called the process *abduction*, *method of hypothesis*, *hypothetical inference*, *method of elimination*, *eliminative induction*, *theoretical inference*, *explanatory inference*, and *explanation-based evidence*. Harman took a stance that this notion was the most basic form of scientific reasoning, and that other inferences, including those from inductive reasoning, could be reduced to this theory.

Forming Hypotheses

An explanation is a proposition of a reason for a particular phenomenon. While there could be unlimited number of explanations for a phenomenon, some philosophers note that a good explanation would minimally be recognized as a *cause* for the phenomenon, be *inferred* from the phenomenon, and be *plausible* compared with other possible explanations. When one is ready to accept the explanation as a likely one, it can be called a *hypothesis* and therefore, in science, is open to testing and rejection.

It is important here to recognize that each of these criteria is evaluated on the basis of likeliness but is also based on background knowledge or theoretical frameworks. The first, *causality*, remains possible only insofar as it can be backed up from inductive reasoning. A cause, after repeated observances, is deemed a cause only once it is seen as sufficiently likely, though not certain, to lead to the effect in question. For every time one drops a bowling ball and it falls to the ground, it is certain only that every time in the past it has been dropped, it has fallen. There is no guarantee that this will happen the next time it is dropped.

The second criterion, *inference*, is also determined on the likelihood that if given the effect, the cause would be inferred. Without inference, there would be two observed phenomena, but they would seem unrelated. It is the relationship to the cause that must be inferred from the effect.

Finally, the explanation must be *plausible*, particularly in relation to other possible explanations. One could say that night and day occur because the sun revolves around the earth, but given the evidence that has been amassed since Galileo, this is not likely and there are better explanations, particularly that the earth revolves around the sun and rotates on its own axis. Often, evidence from many different sources is required to explain an effect better. Some philosophers of science do not think all of these conditions are necessary for an explanation to be deemed *best*.

Reasoning to the Best Explanation: An Example

IBE is a form of reasoning that begins with examining a scientific phenomenon and seeking to understand its cause. Hypotheses are formed, and one by one are eliminated. IBE proposes that all forms of evidence should be considered to form hypotheses of cause and effect, and that the hypothesis which seems best should be accepted. It is important to note that *best* tends to be a subjective determination and is not necessarily the most statistically *likely* explanation inferred from the given evidence. Other forms of inferential reasoning, such as induction, may be employed in IBE. Pure inductive reasoning, however, is limited in that it can only test a hypothesis but is not able to reason for explanations of why effects happen. IBE proposes that enumerative induction is only a special case of inference to the best explanation and is effective only in that it eliminates other hypotheses.

A classic example that is given is the following observation:

(1) The grass is wet.

With this observation, one would immediately begin to seek explanations for why the grass is wet. One would seek to solve the puzzle of what caused the grass to be wet:

Cause → Effect

? → The grass is wet.

First, one may reason that if the grass is wet, then it must have rained. This hypothesis comes from working backward, namely reasoning that when

it rains, the grass gets wet. Therefore, we can see a relationship being formed:

(2) The grass is wet, because it rained.

Or, in cause and effect order:

(3) It rained, therefore the grass is wet.

However, this cause can never be fully endorsed, since only the effect was observed. There may be alternative explanations. Regardless of all of the inductive evidence collected that rain makes the grass wet, it only may remain a possible, perhaps even probable cause. Thus, the explanation becomes the hypothesis of the explanation for the effect.

Say, now, that one was to observe the street, and it is dry:

(4) The grass is wet, but the street is dry.

One now must revise the hypothesis, because if it rained, the street would also be wet. Now one must eliminate the hypothesis developed in (2) and seek for another explanation. One may reason that if the grass is wet and the street is dry, then there must have been a sprinkler system in the yard:

(5) The grass is wet because the sprinkler was on.

Without contrary evidence, this hypothesis explains that the grass is wet better than it rained. So, we may infer from (1) that the sprinkler was on and caused the grass to be wet. We still do not know if this is the cause with complete certainty. It is possible that there was a water balloon fight, a manual watering system, or some other cause. However, this explanation stands as the best explanation until contrary evidence is presented.

IBE in Different Fields

IBE has helped to explain the process of reasoning, especially in science, but also in common sense and other fields. The doctrine of plausibility, particularly, allows scientists to dismiss explanations that analytically or logically may be valid but lack the condition of plausibility. For example, one could claim that the rooster crows and the sun rises, therefore the rooster caused the sun to rise in the sky. While perhaps inductively this could be endorsed beyond reasonable doubt, as the rooster crows every morning and the sun rises every morning, scientists may reject this claim because it simply is not a plausible explanation.

This introduces an element of subjectivity to science, in that no explanation is completely certain and objective. An explanation must remain falsifiable, because it is always open to revision and a possible better explanation may exist. IBE allows inferred explanations to be tested, critiqued, and improved upon. Perhaps most notably, IBE has been used to explain Darwin's process of developing the theory of natural selection. Evidence found in biological diversity does not particularly entail the theory of natural selection outright. Rather, natural selection has been accepted by many to provide the *best* explanation for that evidence.

The theory has applications in other fields as well. Take a criminal trial, for example. The burden of the prosecution is to present evidence, beyond reasonable doubt, that the accused has committed the crime. The prosecution is attempting to argue that their explanation is the best one. If alternative explanations are presented, the burden is to demonstrate that the prosecution's case is stronger and more likely than all other possibilities. Likewise, in history, explanations are proposed and supported by historical evidence. However, they are only accepted insofar as there is no contrary evidence. That there was a true, real human being Napoleon Bonaparte is not provable since we cannot see or talk with him. However, that he was real best explains the plethora of historical documentation and images.

IBE is also related to certain statistical theories, with the influence of Bayesian probability. However, not all supporters of IBE consider the *best* explanation to be the one that is most *likely* or *probable*. Bayesian probability allows for the subjectivity in inferential statistics and also provides evidence to decide which explanation is most likely, or in this framework, the *best*. While Bayesian probability is purely quantitative, IBE allows for a wide range of evidence, but as a principle it is concisely stated in extensions of the Bayesian theorem. For example, take the following:

$$P(H_1|E) = (P(E|H_1)/P(H_1))/(P(E|H_1)P(H_1) + P(E|H_2)P(H_2))$$

This equation expresses that the likelihood of Hypothesis 1 ($P(H_1|E)$) explaining the observed effect is equal to the likelihood that H_1 is a cause for observed E , $\left[\frac{P(E|H_1)}{P(H_1)} \right]$ divided by the likelihood of all, in this case 2, possible causes. Thus, as the likelihood of H_1 increases, the ratio also increases.

We can infer, then, a quantitative measure of how likely H_1 can be inferred from the effect. The equation will reveal the statistical evidence of which explanation is the best. If H_1 is more likely than H_2 , then we would have statistical evidence to support that H_1 is the most likely explanation, though not the only possible one. This Bayesian model allows statisticians to make a defined and quantitatively supported decision of which explanations are most likely. Some critics note that a Bayesian approach offers no rules for accepting a hypothesis but only offers statistical probability. IBE, at some point, does set a rule to accept the best explanation.

Critiques

IBE has been critiqued for its subjective epistemology. Karl Popper (1902–1994), a renowned philosopher, observed that IBE allows one to declare something as possibly or probably true. According to IBE, once an explanation has satisfied the scientist as the most likely to be true, it can be accepted, essentially, as true. However, critics push for a less subjective approach in science and require all hypotheses to be declared not as an expression of truth but rather as, simply, not yet refuted. Even though evidence may support one explanation more than other, the terminology of one being the *best* introduces a high level of subjectivity. Thus, IBE, though opening the possibility for an explanation to be refuted, allows scientists to declare their explanation as *best* (even determinately so) in the absence of what the scientist deems to be a better one. Additionally, IBE has been critiqued for being merely a type of inductive reasoning.

Conclusion

When seeking to explain phenomena, one infers possible causes. These explanations are then evaluated and the explanation that seems the best explanation, given all evidence, is accepted. Though one may never know a true cause for an effect with complete certainty, IBE maintains that the best explanation should be endorsed. Thus, the process of inferring explanations involves reasoning from phenomena to an explanation for it.

Joseph C. Rumenapp

See also Hypothesis Testing; Hypothetico-Deductivism; Inferentialism; Refutation; Statistical Inference, Bayesian

Further Readings

- Harman, G. H. (1965). The inference to the best explanation. *Philosophical Review*, 74(1).
- Josephson, J. R., & Josephson, S. G. (1994). *Abductive inference: Computation, philosophy, technology*. Cambridge University Press.
- Lipton, P. (2004). *Inference to the best explanation*. Routledge/Taylor and Francis Group.
- Okasha, S. (2002). *Philosophy of science: A very short introduction* (Vol. 67). Oxford University Press.
- Okasha, S. (2000). Van Fraassen's critique of inference to the best explanation. *Studies in History and Philosophy of Science Part A*, 31(4), 691–710. [https://doi.org/10.1016/S0039-3681\(00\)00016-9](https://doi.org/10.1016/S0039-3681(00)00016-9)
- Psillos, S. (2004). Inference to the best explanation and Bayesianism. In F. Stradler (Ed.), *Induction and deduction in the sciences* (pp. 83–91). Springer.

INFERENTIALISM

Inferentialism emerged as a theory of semantics to rival a once-dominant view that language is primarily representational. Representationalism, the long-standing view that linguistic signs represent objects in the world, was largely a common-sense understanding of language that dominated philosophy and science throughout much of Western history. Inferentialism, however, is a framework built around the notion that the meaning of a linguistic expression is found in its relationship to other expressions. Inferentialism proposes that (a) the act of being meaningful is ultimately an act of making connections between assertions (inferring) and (b) that these acts are governed by rules. The basis of inferentialism rests on the pragmatic notion that meaning is an act of doing. The primary way of doing meaning is through giving and asking for reasons, and therefore, it has often been referred to as a type of *linguistic rationalism*.

Representationalist theories of semantics underlie many of the philosophies in modernist thought, taking as a fundamental tenet that words, and other signs by extension, in some way *stand for* objects in the world. This representationalist

family of theories supposes that speakers use language to represent the world so that receptors of communication could thereby understand meaning contained in the signs. Inferentialism, by contrast, is a theory of meaning, suggesting that one interprets or *infers* meaning. This view has been applied to logic, philosophy, and semantics. Semantics, in this sense, is not solely about the word as such but rather how that word is used. Thus, inferentialism resonates well with American pragmatism. Many pragmatist thinkers point out that the focus on *what meaning is*, and *how things are meaningful*, should not be primarily understood at the word level but rather at the sentence level. That is, they proposed that semantics answers to pragmatics. Later thinkers have more directly developed this line of thinking to suggest that meaning is not represented by signs but rather is inferred from the use of signs.

While inferentialism may be used to theorize about semantics in general, it has special applications in logical and philosophical language (e.g., the logical connectors). The meaning of logical constants, for instance, can be understood as inferentially derived, thus allowing for forms of proof-theoretic semantics to emerge. In what follows, the history and fundamental tenets of inferentialism are considered; then applications to logical notation will be reviewed; and finally, a broader framework for semantics is presented.

Historical Foundations

Robert Brandom most recently and thoroughly codified the theory in his book, *Making It Explicit*, explaining that language makes explicit otherwise tacit relationships that are inferred. The term *inferentialism*, though coined by Brandom, can possibly be traced back to Immanuel Kant (1724–1804) and, more directly, to the language philosophers of the 20th century. Tracing the roots of inferentialism to Kant's transcendental idealism allows inferentialists to reject both pure rationalism and pure empiricism. At once, they deny knowledge is a priori, or that there is a world *out there* which we can know with no prior experience, while also denying that the mind is simply a blank slate. Specifically, in inferential approaches to epistemology, language is the focus of analysis. Knowledge is thought to reside apart from individual human experiences, and therefore

it is considered anti-empiricist. However, it is important to understand inferentialism not only as taking on the formal practices of reasoning, but that it is made explicit through language. Therefore, it does not follow the rationalist paradigm that knowledge is independent from human practice and is merely reasoned. Thus, the paradigm representative of an inferentialist epistemology can be traced at least as far back as Kant.

Another important step in the development of inferentialism was the influence of Ludwig Wittgenstein (1889–1951). In his later years, he proposed the notion that language was a kind of game. That is, the words were to be understood as functioning similarly to pieces in chess, and they had roles similar to the types of pieces. Hereby, meaning and language were linked, through the *use* of words and the *role* they play. Thus, language is rule-governed in the sense that it is used to play different games. Wittgenstein posited that there could be any number of language games, and he did not privilege any of them in his theory of language. While most of his foundational works were published posthumously from his lectures, his movement was toward viewing meaning as found in the use of a sign within a particular game, rather than in the sign itself. This marked a major shift in the way language was conceptualized.

Another proponent of this paradigm was Wilfrid Sellars (1912–1989), an American philosopher of the mind. Sellars, like Wittgenstein, saw semantics as studying the content of words in relation to their use. So, semantics was subjugated to rules of use rather than intrinsic content. Sellars noted that there was a type of rule that was, perhaps, more specifically responsible for meaning, namely, inferential rules, which was dubbed *the game of giving and asking for reasons*.

These theories focused on the pragmatic use of language, indicating that language does not merely *reference* things outside of a given context but rather is used in particular moments to *do* something in the context. Therefore, the meanings of words were not derived from their relationship to an object in the world but rather their use or performance on the world.

Brandom took a stance that the game of inferring is the primary game humans play with language, thus diverging from other postmodern followers of Wittgenstein. He proposed that *assertions* were the primary speech act of language. For

something to qualify as a linguistic practice, that is, giving and asking for reasons for these assertions, one was disposed to acknowledge the normative statuses of commitments and entitlements. Thus, the appeal to normative statuses of commitments and entitlements allows inferentialism to claim a sense of objectivity while maintaining those statuses as socially developed. In this way, inferentialism maintains that it is *linguistic rationalism* and also pragmatic in nature.

Key Concepts

Inferentialism can be seen as an expressive, ongoing process, or game, of being meaningful. One can look at language, then, as largely being concerned with linguistic moves that commit people to certain claims. Commitments and entitlements are key concepts to understanding inferentialism. When one makes a claim, or an assertion, there are always tacit, rational questions of *what that claim commits the speaker to* and *what is necessarily endorsed in order to be entitled to make the claim*. Thus, when one asserts, one may be liable to justify that assertion or give reason for it. So, when one commits to a claim, she opens up the possibility that another may demand justification for that claim. Any linguistic movement, then, not only adds commitments to assertions but also opens up the need to give reason for being entitled to that claim.

Commitment and entitlement compose the two major normative rules of the language game of inference. One must always *infer* to all those possible consequential commitments entailed by an original statement, and one must always *infer* the possible reasons for the statement to which they were entitled in the first place. The metaphor of gaming and scoring, then, is that when given an assertion, one keeps a score of possible future statements to which one could also be committed and possible previous statements that serve as entailments to a particular claim. The *game* is broadly a pragmatic theory, one of *language in use*. However, the pragmatic game can be stretched to become a semantic theory as well, that is concerned not only with use of sentences but also with content of words.

For instance, the *game of giving and asking for reasons*, as Brandom explains, hinges on the idea of keeping score of commitments and entitlements.

As one makes an assertion, the person is thereby committed to that concept. Also, the person ought to be committed to other assertions as well. Brandom uses the following example:

“The swatch is red.”

By using this statement, as a move in a game, one adds to the score that the speaker is thereby committed to the fact that the swatch is red. Additionally, the following statement ought also to be added to one’s score.

“The swatch is colored.”

The commitment to *red* thereby also commits one to *colored*. This could be called a *consequential commitment* because a commitment is made by virtue of being a consequence of an already explicit assertion. Each linguistic move must be treated as a commitment; otherwise, it has no real effect and thus is not rational. Herein, we can see how inferentialism receives its name. It is the relationship between concepts where meaning is inferred. The very rule that one who commits to “The swatch is red” is obligated to also commit to “The swatch is colored” demonstrates the inferential relationship between the linguistic sign and the conceptual content of the signs.

The key concepts of commitments and entitlements allow for inferentialism to be expressed as a form of linguistic rationalism by the likes of Brandom, to attempt to achieve a sense of perspectival objectivity. It also allows inferentialism to take up specific roles in logical and philosophical language. In logic and philosophy, there are often words that are considered constant in their meaning. Connectors, for example, have a very specific semantic role and a very restricted meaning. This allows for specificity in logical notation. Thus, applications of inferentialism in logic are considered.

Linguistic Rationalism

Inferentialism, at least in some varieties, seeks to ensure a particular type of objectivity, particularly through establishing commitments and entitlements as norm-governed acts. By *objectivity*, it should be clarified that what is at stake is not the truth of claims and whether they directly correspond to reality but rather whether the reasoning

through language is valid. This is the burden of the practitioner and is always open to further appraisal.

Objectivity, in the sense of matching the real world, is not possible and is not even a question for inferentialism. Indeed, when dealing with conceptual contents, one of the major critiques of inferentialism by representationalists, is that there seems to be no objective meaning. Some forms of inferentialism attempt to deal with this issue by proposing that objectivity exists from the view of a speaker. One makes claims that are rational and because they can be countered or reinforced. They also are objectives in that speaker's view. Because the rules for inference are normative, it is possible for people to speak to one another, viewing these practices as objective in nature, while they are in all actuality not objective. This type of linguistic rationalism seeks to secure the possibility of objectivity of thought, making the theory of inferentialism especially applicable to logic, philosophy, and even the sciences.

Beginning with the premise that any linguistic community is a rational linguistic community, Brandom claims that the following two statements differ significantly from one another:

- (1) The swatch is red.
- (2) It is assertable by me that the swatch is red.

Sentence (1) does not indicate *who* is committed or entitled to the claim, and thus it is possible to extend the claim to all rational beings, while Sentence (2), on the other hand, specifies *who* is committed and entitled to the claim. The two, though broadly similar, differ in the sense that they are incompatible under certain conditions. In both, the speaker is committed to the assertion; however, the entitlements are indeed different. The first requires investigation of the color of the swatch, whereas the second merely requires the investigation that the speaker has made the assertion. This difference, for Brandom, is the key to allowing for objectivity of thought. The *normative fine structure of rationality*, as he calls it, allows one to make objective assertions, following the normative reasoning of linguistic rationalism, despite social aspects of claim making and individual attitudes. Thus, the question of any assertion in relation to its propositional truth is whether the

practitioner has followed the normative rules of the assertion game.

In this game of giving and asking for reasons, there are three inferential roles that emerge. The first, *commitment preserving*, generalizes sentences that are used deductively. But deducing from previous assertions, it thereby preserves the commitments that are made. *Entitlement preserving* statements, generally, are inductive in nature. They amass support for a claim and thereby lend evidence for the speaker's commitment to an assertion. Finally, *incompatibility relationships* are inferential roles a sentence plays in commitment to incompatible claims. For example, "the car is red" also commits the speaker to "the car is not blue." These inferential roles are not the acts of inferring as such, but rather the marks on a *score-card* that are being checked off as commitments are made and consequential commitments are also marked off. Essentially, then, these are actions that occur to make everyday communication possible, as we tick off the *score* of others. We note the implication of the claim and the evidence laid for the claim. In this process, the normative rules of inferring from sentence to sentence allow us to understand the rational movements of the other.

Inferentialism, as can be seen, is a linguistic rationalist perspective on truth and relies on deductive reasoning between assertions that are made as well as those to which one is committed via the consequences of the assertion. Humans deduce meaning from these assertions. However, unlike rationalism proper, inferentialism does not begin with the premise that the world has an autonomous rationalist structure that humans can deduce outright. Inferentialism, rather, focuses on the innate, rational ability for humans to deduce, through reason, the relationships between premises, which may or may not correspond to the world-as-it-is.

Applications in Logic

Inferentialism has special applications to logic. Since the paradigm gives up the representational quality of signs in favor of an inferential quality, the logical vocabulary can be seen as a tool to make the implicit inferences between nonlogical commitments explicit. Particularly due to its claim of objectivity of mind, it can be applied to logic proper.

Specifically, inferentialism has implications for logical expressivism, that is, that logic is not reducible to mathematically inferential symbols but rather is intended to be used with real content.

Gottlob Frege (1848–1925), a German mathematician and one of the founders of modern logic, opened up the possibility to view logical connectors as governed by rules of proof, and therefore expressive in nature. Logic, in this view, then, is not simply a set of logical connectors but centers on the logical connectors as they make inferential commitments explicit. Logic, in this model, can be understood as the process by which the implicit relationships between content is explicitly stated.

Logical vocabulary, according to inferentialism, is derived from the use of inference in logical patterns. Inferentialism would hold the following, basic, logical test:

Abstract, logical example:

$$\begin{aligned} P \& Q &\rightarrow P \& Q \\ P \& Q &\rightarrow P \\ P \& Q &\rightarrow Q \end{aligned}$$

Specific, concrete example:

The sun is yellow. The sea is blue. The sun is yellow and the sea is blue.

The sun is yellow and the sea is blue. The sun is yellow.

The sun is yellow and the sea is blue. The sea is blue.

However, inferentialism holds to the meaning of $\&$ and $\rightarrow$ not because it is an autonomous system or because the world is explained through these logical connectors. Rather, inferentialism holds this test to be true because of the pattern of use that constitutes the meanings of the constants. The *conceptual content* of the logical signs can be inferred from their use. One may critique, as Arthur Prior, the logician, has, that this approach could allow for the existence of a logical contradiction and thus is a flawed approach. You cannot simply use a word “tonk” and infer meaning. The logic behind the statements would still be flawed. The simplified example that follows demonstrates that inferential rules alone do not constitute meaning.

Abstract, logical example:

$$\begin{aligned} P &\rightarrow P \text{ tonk } Q \\ P \text{ tonk } Q &\rightarrow Q \end{aligned}$$

Specific, concrete example:

The sun is yellow. The sun is yellow tonk the sea is blue.

The sun is yellow tonk the sea is blue. The sea is blue.

One cannot reason from the “sun is yellow” to “the sea is blue” outright. Even given the first example in which “tonk” connects the “sun is yellow” to “the sea is blue,” the inference could not be made that “the sea is blue,” because “tonk” fails to fix the meaning. “Tonk,” up until now, has no logical system in which to exist. Upon the second example, however, one could infer “tonk” as a logical connector (in this case with the semantic range of “and”) because it passes the logical test. Therefore, all logical connectors must be reasoned from their use. Critics propose that it is not possible to understand meaning as being conferred through normative patterns of use. However, that the tonk example exists does not necessitate that “tonk” is indeed a true logical connector in inferentialism. Logical constants are inferentially derived. Therefore, this is a semantic view of logic in which the language of logic abides by the concepts of use.

Inferentialism laid the foundation for *proof-theoretical semantics*, in which the reasoning between truths is ultimately an inferential activity. Proof-theory had long allowed for the logical deduction of truths from premises but only insofar as the logical syntax would allow. This is most clearly seen in mathematics in which formal proofs are treated as mathematical objects. The syntax allows for the proofs to undergo mathematical functions, following the rules of deduction. The study of the semantic component of logic was an innovative shift, leading away from resting epistemology on truths derived abstractly and syntactically but rather on the justifications of the claims derived semantically. We see this clearly in mathematics, as well, in which models of theories are the objects of study (not the proofs themselves). Thus, proof-theoretical semantics is an investigation into the knowing-how logical deduction is done.

Proof-theoretical semantics has allowed the inquiry into rules that govern logic. This has applications in mathematics and logic, as the inferential reasoning between concepts allows for the normalizability of proofs. Inferentialism is not concerned with the logical constants as representational figures but rather as the knowing-how to express already implicit relationships between assertions. As can be seen, inferentialism is a framework that is used in logic and philosophy. However, it has broader applications in semantics.

Semantic Theory

One variety of inferentialism is *conceptual role semantics*. Conceptual role semantics proposes that meaning can be seen as *causal*, based on individuals' reasoning or deducing through language from one assertion to another. This is different from some other varieties, which propose that meaning is *normative*, that is governed by rules of inference. In both types of theories, the act of inferring is primary, however, *how* that act occurs may differ. Conceptual role semantics, as a subset of inferentialism, helps to understand the realm of semantics. A related theory, *inferential role semantics*, arose in philosophy and is somewhat more restrictive to cognitive roles signs play in inference. The focus in these theories is what constitutes the meanings of linguistic expressions. Diverging from theories of semantics that assume a representational relationship between expressions and the world, in this paradigm the concern is with the role an expression plays in one's cognition. Thus, we can see that the relationship of interest is not the word and an object in the world, but rather the use of a word in linguistic expressions. This "use theory" of semantics generally falls under the domain of pragmatics, then. The meaning of words, or *semantic* meaning, is subject to their *pragmatic* use.

In linguistic theories, however, the distinction between semantics and pragmatics is an important analytical one. Here we may look at the relationship between the two, anchoring the discussion with considerations in syntax. Syntactic forms, such as declarative sentences, on the surface level, have no meaning independent from concepts and use. Thus, a sentence like the example introduced

earlier, "The swatch is red," is declarative in form. However, apart from the meaning of the concepts and the context of use, the sentence alone is meaningless. Likewise, once we consider the semantic domain, we will see that there are concepts involved. The concept of *swatch* and the concept of *red* are important to the meaning of the declarative sentence. However, there may be dozens of types of swatches and dozens of shades of red. Yet, we know we can talk about the concept of *swatchness* and the concept of *redness*. The commonality between all swatches and all reds is what we may call *conceptual contents* or *propositional* contents. Yet though there is some sort of content applied to the signs, the meaning is still unknown until we take into consideration that this declarative sentence is actually doing something. It is being used in a particular way, namely, to make an assertion. This is the pragmatic use of the expression. We can see, then, how linguistic expressions rely on syntactical forms, semantic conceptual contents, and pragmatic roles.

Inferentialism takes up this logic by claiming that the link to meaning is made in the transition between linguistic moves (some spoken, others unspoken but entailed). Semantics, in this way, will always fall short of explaining meaning unless the pragmatic use is considered. "The swatch is red," is, of course, only meaningful in relation to external information, or the relationship to other committed stances ("The swatch is colored") and from entitled stances (that one can critique the assertion to ask for reasons).

Within these *use theories* of language, it may be tempting to get caught in an endless cycle of meaning deference. That is, if a word, or any linguistic sign, has no intrinsic meaning and is arbitrary, then one may seek ad infinitum for a cause or explainer. However, inferentialism does not seek to define words in this manner but rather looks to the word's role in inference. Herein lies the significance of a word.

Conceptual role semantics is a theory of meaning but not a theory of truth. This is an important distinction because some philosophers have advocated for a *truth-conditional* view of semantics. In these theories, truth and meaning are tightly bound. The truth of a linguistic expression then

would depend on the truth of the parts, or the conditions, which would make the statement true. Conceptual role theories privilege the role an expression plays in a speaker's mind, and since no two speakers have an identical psychological state, the concept will be different, albeit slightly, perhaps. The concepts serve as a tool to allow inferences to be made and meaning is roughly found in the role the words play, not in the relationship to external conditions of truth. Meaning, both in semantics and in logic, is found in the inferential role signs play, not the representational or conditional truth of the sign itself.

Variations of Inferentialism

A variety of forms of inferentialism exist. Brandom's most widely articulated variety stands as a strong form, that is, the inferential use of language is primary among all potential uses, as noted earlier. This premise extends to any type of linguistic practice. Weaker or more moderate forms tend to apply this premise only to logical vocabulary, or that which is used to reason to and from set assertions. Michael Dummett and Dag Prawitz, also identified as inferentialists, have a much more restricted claim of language's inferential function in that it applies to logic specifically. Brandom's inferentialism expands upon these inferentialist approaches and thus is able to apply the "-ism" to an ongoing movement.

There are also varieties of inferentialism which employ inferences in the wide sense, that is, there are many types of inferences which may count in linguistic practices. Alternative views may restrict the inference merely to the operators in language. We begin to see, then, that inferentialism most clearly has its applications in logic, with some theoreticians extending its use beyond logic to the whole of language. However, there also is a variation on how inferentialist approaches relate to logic. Niel Tennant, for example, carries the Dummett and Prawitz trend to apply inferences within an intuitionistic logic, thereby reforming classical logic and taking an anti-realist position. This stands in contrast to Brandom's use of inference in classical approaches to logic. The variations, then, can be seen to exist on the types of inferences allowed and the application of the inferentialist approach to language in general or logic in particular.

Conclusion

Inferentialism, as we have seen, is a paradigm in the theory of meaning. It has historical roots in Kantian philosophies but was developed more thoroughly by logicians. In logic, inferentialism explicates the expressive function and inferential nature of logical connectors. Logical connections are not seen to exist autonomously but are rather ways of making inferences explicit. Mathematical and logical notation, for example, can be inferentially understood as expressing the relationship between content. This allows for proof-theoretical semantics to study the language and categories of logical notation, similar to model theory in mathematics. Thus, it is a semantic view of logic as opposed to a syntactic view. The formal connections are only meaningful if they are materially inferential as well. That is, formal rules of logic and rules of inference are not synonymous. Inferences must be materially good, that is, the content of the concepts must be inferentially related, thus deriving the formal rules of inference.

In the broader program of inferentialism, language as a whole, and even human cognition, is considered. In some varieties of inferentialist semantics, such as conceptual role semantics and inferential role semantics, the emphasis is placed on the role concepts play in human cognition and how meaning is inferentially derived from concepts. Other forms focus rather on the normative roles that concepts play in inference. In this later view, the rules of inference are seen to be normatively governed, that is, socially governed by way of giving and asking for reasons. The primary function of language, in this view, is to make assertions that commit the speaker to the conceptual content of the claim, and consequentially to other conceptual content as well. By making a commitment, however, one also is opening up the opportunity for the entitlement to that commitment to be questioned. Thus, every linguistic move shifts the scorecard, so to speak, in this language game.

Joseph C. Rumenapp

See also Intuitionism in Logic and Mathematics; Logic, Formal and Informal; Logical Theory; Theories, Semantic Conception of

Further Readings

- Brandom, R. (1994). *Making it explicit: Reasoning, representing, discursive commitment*. Harvard University Press.
- Brandom, R. (2000). *Articulating reasons: An introduction to inferentialism*. Harvard University Press.
- Dummett, M. (1993). *The seas of language*. Clarendon.
- Frege, G. (1971). *Begriffsschrift* [Concept-script]. Verlag von Louis Nebert. (Original work published 1879)
- Peregrin, J. (2012). What is inferentialism? In L. Gurova (Ed.), *Inference, consequence, and meaning: Perspectives on inferentialism* (pp. 3–16). Cambridge Scholars Publishing.
- Prawitz, D. (2006). Meaning approached via proofs. *Synthese*, 148. <https://philpapers.org/go.pl?id=PRAMAV-2&proxyId=&u=http%3A%2F%2Fdx.doi.org%2F10.1007%2Fs11229-004-6295-2>
- Prior, A. (1960). The runabout inference ticket. *Analysis*, 21.
- Sellars, W. (1953). Inference and meaning. *Mind*, 62.
- Tennant, N. (2009). Inferentialism, logicism, harmony, and a counterpoint. In A. Miller (Ed.), *Essays for Crispin Wright: Logic, language, and mathematics*. Oxford University Press.
- Wittgenstein, L. (1953). *Philosophical investigations* (G. E. M. Anscombe, Trans.). Blackwell.

INFORMATICS

Informatics is a loosely defined field of study generally focused on the creation, acquisition, transformation, and application of data. In practical use, there are meaningful distinctions between the academic field of informatics and the commercial implementation of related concepts. The term does not have a single unifying definition or agreed-upon framework the way some other fields may. Depending on the context and goals of a scholar or practitioner, the term may refer to work in the areas of data analysis (mathematics), algorithms (computer science), physical information exchange systems (electrical or network engineering), or even the public dissemination of data-driven knowledge (communication studies). Contemporary usage generally means any work labeled as informatics will be centrally focused on computer technology. Because the field is relatively new and is interpreted in different ways, this

entry begins with its etymological origins as a way of demonstrating the range of work reflected in the range of theories employed. Broader conceptual frameworks are addressed, followed by individual theories looking at technical and human aspects of the field.

Etymology

Because there is not a single, widely used definition of informatics, it is useful to know the history of the term as the evolution of the moniker reflects rapid changes in the field itself. Informatics is a relatively young field compared with other science, technology, engineering, and math (STEM) fields. Because of this, there are numerous connections to other STEM disciplines, but formal organization of work in the area is more recent and less rigorously organized.

The term *informatics*, in the English-speaking world, is sometimes assumed to be derived from the combination of “information” and Latin-derived “mathematics.” However, it is more likely to be a translational carryover from the German term *informatik* applied to the concept of data processing in the 1950s. Similar constructions follow such as the French *informatique*, Spanish *informatica*, or Russian *информатика* (informatika). Regardless of the specific language, the term seems to generally have a broad, inclusive definition regarding computing and data when in popular usage.

Conceptual Approaches to the Field

In more scholarly contexts, the varying definitions of informatics are overlapping but can be quite distinct depending on the source or application. Informatics Europe focuses primarily on the scholarly pursuit of the field and represents universities, research labs, and their partners based in Europe. Founded in 2005, the group closely identifies the field of informatics with areas often titled computer science, computer engineering, information technology, or information and communications technology. Conceptually, their usage may also refer to the study of human–computer interaction (HCI). Usage of the term across Europe is more likely to refer to these types of programs. Therefore, a European university is less likely to have a

department of informatics and a department of computer science, as they would appear redundant. In this view, informatics faculty or students may be developing specific theories of software development or designing new computer hardware in order to fill computational needs. This often requires strong collaborations with mathematics, physics, or chemistry.

In the United States, the term often has a slightly different connotation focusing more specifically on data and how they are created, stored, transmitted, or applied. For instance, some degree programs are interdisciplinary, combining mathematics, computer science, and other fields in which the work can be applied (e.g., biology, political science, journalism) rather than a stand-alone department. In other cases, institutions have departments or schools of information containing multiple degrees or programs of study each focusing on different aspects of the overall concept of informatics. Scholarly fields that focus on creating technologies and tools are, generally speaking, older and more established than the more recent higher learning approaches to studying data itself. Consequently, there are many more academic programs that identify as computer science or electrical engineering in the United States than there are those that refer to themselves as informatics. A U.S. American university is more likely to have both a department of informatics (or information science) and a department of computer science, as they would occupy different spaces in the overall technology landscape of the organization. In this view, informatics faculty or students may be acquiring and investigating data as a focus, only employing software and computing systems as a secondary need to fill their primary mission.

Schools of information in American universities may also conceptualize informatics more as the systematic study of how to organize, archive, and make accessible to the public knowledge derived from data. For instance, the term *data science* refers to many of the same efforts and conceptual work as informatics but more explicitly focuses on data itself rather than the systems that make use of data. This might be conceptualized as a field directly related to informatics or as a subdomain of informatics itself, depending on the focus and design of the institution. For example, computer science may develop techniques for natural

language processing, while informatics or data sciences seek to develop frameworks for creating and mining large corpuses of human utterances. Similarly, an informatics approach to artificial intelligence might focus on the large data sets necessary for machine learning but not on developing the models themselves.

There is a global trend across organizations that train professional librarians to either align with informatics and similar programs or outright combine them into a single approach to information management in a field sometimes referred to as *library and information science*. The American Library Association aims in its mission to promote library and information services. The Chinese government combined programs in their national attempt to promote automation technology in libraries into a single digital library and information systems approach. Support of these processes and public services requires supporting or relying on work common in the field of informatics.

In more commercial or applied contexts, the term *informatics* tends to be strongly associated with collecting large amounts of data from individuals and groups in order to identify trends, make predictions, and develop products or services. Though the utility of analyzing data is not new, since the start of the 21st century, analytical analysis of information has rapidly become a fundamental component of the business, nonprofit, and government/social service sectors. Of particular note, marketing and advertising industries have seen heavy reliance on developing methods of collecting detailed information not just on large audiences but also on individual people. Data about their demographics, past experiences, and current activities are of potential use in making predictions about their future activity. This is valuable to a business marketing a product or a social service entity attempting to change the activities of people in the public.

Health care is an area of particular importance regarding analytics because there are a number of stakeholders and potential uses for collected data. For example, a hospital may track multiple factors about each patient receiving treatment. These data can be analyzed to evaluate hospital and doctor performance while seeking to improve outcomes or evaluate distribution of limited resources. However, the same data might be useful to

companies developing and selling medications or health insurance. Governments and related social organizations might use this data to support public health campaigns or design legislation. In turn, individuals may be motivated to use digital tools and wearables to improve their own health, such as popular health and activity trackers. Each of these vectors is intertwined in complex ways that can support business and health outcomes, but only if there are ways of collecting, processing, and making use of the data. Informatics plays a key role in these kinds of endeavors whether it takes the form of designing health-measuring electronics, analyzing data, or supporting recommendations for health strategies.

Theoretical Constructions

Because informatics is a new field relative to many of the ones it draws from or connects to, many of the theoretical constructions used in scholarship or practice often come from other places. The term *debtor field* refers to an area of study that primarily draws from others in order to execute its work. If considered from the viewpoint that informatics is essentially another term for computer science, network engineering, or information technology, then most of the theory employed naturally comes from those constituent fields. Alternatively, if informatics is viewed as its own relatively new field that looks at information science but connects to other STEM fields, much of the theory used may be labeled as coming from other fields, and informatics itself may seem theoretically sparse.

In any case, the theoretical concepts and approaches used in informatics can be most easily categorized based on the goal of the work. The fundamental purpose of a theory is to explain. In informatics, theories generally focus on large-scale, epistemological aspects of finding and using data (conceptual frameworks) or on narrowly tailored approaches to understanding how a specific problem or question is addressed (theories).

Conceptual Frameworks

Systems theory is an approach to understanding people and technologies by treating them as systems or machines. Humans are generally capable

of considering how a system works through chronological or cause-and-effect relationships. Systems have interconnected parts that are dependent on one another to complete a large task that no single part could complete on its own. Combustion engines, desktop computers, and factories are examples of systems. The systems approach takes a complex environment of interacting components and studies it by using many of the same techniques a mechanic would use on an engine or an engineer would use in a factory. Looking for the order in which events occur, what effects they cause, and how small changes can significantly affect system outcomes allows a researcher or practitioner to identify which technologies or people might be changed in order to produce a desired outcome. Systems theory can be used to study human-made technologies, human communication, or even naturally occurring phenomena.

Related to systems theory is another framework called *cybernetics*. Rather than considering a single, self-contained system, this approach assesses the interaction and influence of intelligent systems on one another. The theory assumes that a system is intelligent insofar as it has a goal of some type and reacts to outside stimuli to continue achieving that goal. Depending on the specific application of the theory, it can be used to study how systems handle information processing, feedback, and control in addition to how they can regulate each other or respond to interruption. This framework has been applied to a wide range of technologies and human communication.

Humans typically have definable purposes for using technology. *Uses and gratifications theory* has been used in a number of scholarly fields since the 1940s. Originally focused on understanding why people select and use the communication media available to them, it has been significantly expanded and adapted since that time. This approach is no longer considered as strong as others by a number of scholars, but the essential concept that people do things to satisfy their needs has underpinned a number of other important theories with regard to humans using technology. A related theory, *activity theory*, is a more socio-cultural approach to understanding what motivates people and how they relate to the objects and processes necessary to accomplish their goals.

Theories

Because the field of informatics does not have a clear boundary and is instead still developing and overlaps with others, there is a broad range of theories applicable to the field, reflecting the wide array of questions it attempts to answer. A brief selection of useful theories is presented here.

From the perspective of computer science, mathematically driven theories are particularly useful. *Analytic number theory*, broadly conceived, is a mathematics theory that compares the outcome of a function with its approximation. A more specific use of this in computer science is in applying mathematical functions reflecting the design of software to make predictions about how complex specific algorithms are in order to determine their efficiency and optimal use. In computer science and informatics, the most common use of this approach is referred to as *Big O notation*. *Graph theory* looks at relationships between paired items in a network of related components. It provides mechanisms for analyzing and optimizing the connections in a system to better find relationships between data points. For example, software to generate driving directions might treat cities as nodes paired (related to each other) by the roads that connect them. Routing equipment uses graphs to compute best paths for data to take through a network.

In many cases, it is preferable to have a general device that can solve many problems rather than a specific machine for each problem. Alan Turing's work in the early days of computing contributed immensely toward this end. A *Turing machine* is a device that solves a range of generalized problems based on specific input and rules. Essentially, theories around Turing machines have to do with calculating whether a given machine will be able to solve a particular problem within specific constraints. The ability to mathematically calculate whether a system can solve a problem is incredibly valuable in designing both hardware and software. Computer scientists, computer engineers, and mathematicians all work together in this area, designing technologies capable of reliably solving problems and directly or indirectly proving they will work rather than investing in a system that will ultimately fail.

Moving away from the purely mathematical and toward the interaction of machines and people, a number of theories and approaches are useful. Originally drawing from psychology, a *theory of affordances* describes qualities of an object that suggest how the object should be used by people. The idea is that designing an object so that it is obvious how it should be used is more efficient and intuitive than a more abstract object that requires labels or instructions. Examples such as a hole the same shape as a peg, a knob that only turns in one direction, or lights meaning something is turned on make the interface for a system easier without training. Different versions of this approach disagree over whether the affordances should or should not be culturally sensitive or whether perceived affordances are the same as objectively true ones. In any case, the theory helps explain how objects should be designed so that people can make most effective use of them.

In the 1990s, *information-foraging theory* was developed by treating how animals seek out food as analogous to how humans seek out information. It proposes that people make rational decisions based on the perceived value of an information source relative to the difficulty in using it. For instance, a website that contains a large amount of information in an easily accessible format is a high-value source relative to a book that contains less information and lacks the ability to readily locate specific topics. One could use this theory to predict the success of a source such as Wikipedia over traditional print encyclopedias. Information foraging is conceptually related to the psychologically based communication concept known as *social exchange theory*, which posits that people manage their relationships by evaluating the return they receive for an investment of time and information disclosure. Information-foraging theory is useful for predicting the types of sources people will use and how best to present data for particular audiences based on their abilities and motivations.

Gordon S. Carlson

See also Artificial Intelligence; Computer Science; Cybernetics, 20th Century; Framework; Mathematics, 20th Century; Mental Models; Software Engineering

Further Readings

- Bertelsen, O. W., & Bødker, S. (2003). Activity theory. In J. M. Carroll (Ed.), *HCI models, theories, and frameworks: Toward an interdisciplinary science* (pp. 291–324). Morgan Kaufman.
- Chartrand, G. (1985). *Introductory graph theory*. Dover.
- Gibson, J. J. (1979). *The ecological approach to visual perception*. Houghton Mifflin.
- Informatics Europe—Home. (n.d.). Retrieved May 25, 2022, from <https://www.informatics-europe.org/>
- Katz, E., Blumler, J. G., & Gurevitch, M. (1973). Uses and gratifications research. *The Public Opinion Quarterly*, 37(4), 509–523. <https://doi.org/10.1086/268109>
- Pirolli, P., & Card, S. (1999). Information foraging. *Psychological Review*, 106, 643–675. <https://doi.org/10.1037/0033-295X.106.4.643>
- Turing, A. M. (1996). Intelligent machinery, a heretical theory. *Philosophia Mathematica*, 4(3), 256–260. <https://doi.org/10.1093/phimat/4.3.256>
- von Bertalanffy, L. (1976). *General system theory: Foundations, development, applications* (Rev. ed.). George Braziller. (Original work published 1968)
- Wiener, N. (1948). *Cybernetics: Or control and communication in the animal and the machine*. MIT Press.

INFORMATION THEORY

Information theory is the name given to the study of signaling and communication as an abstract field, separated both from the physical problems of building communication systems and from the meaning of the information communicated. The theory is principally concerned with operationally well-defined ways of quantifying uncertainty: Its focus is upon the probability of potential signals being sent. The fundamental unit of this quantification is a *bit*, which represents complete uncertainty regarding which out of two possibilities is the case. The Shannon information measure generalizes this to an arbitrary number of possibilities and to varying degrees of certainty.

The two most important applications of information theory are to data compression, the efficient and accurate storage of data, and to error correction, the efficient transmission of data

without errors introduced by noise. Information theory underpins almost all modern information processing and communication technology. The following sections explain the relationship between information and uncertainty, in general and in situations where a correlation exists between different systems. Next, the operational properties of the Shannon information measure are described, and the entry concludes with a discussion of the relationship between information theory and extant notions of computational complexity.

Information and Uncertainty

The Shannon information measure was introduced by Claude Shannon in 1948 as a means of quantifying the uncertainty of a source of signals but has been generalized to uncertainty about any general state of affairs.

Communication systems involve encoding messages into a standardized set of signals, which are transmitted from source to receiver, and then reconstructing the sent message from the received sequence of signals. Simple examples of this include Morse code and semaphore. By focusing on the abstract properties of such coding schemes, Shannon was able to identify key properties of information common to all such systems.

Each possible message has a particular probability of occurring for a given signal source. The Shannon information of the signal source is given by the expectation value of the logarithm of these probabilities. It has a simple operational interpretation: To find out which message was sent by asking a sequence of yes–no questions, the best possible strategy will, on average, require a number of questions equal to the Shannon information. This strategy is equivalent to coding the message as a sequence of ones and zeroes, which represent the yes or no answers to the optimal sequence of questions. The ones and zeroes of the coding scheme are binary digits, or bits, and the Shannon information measure represents the average number of bits required to encode messages from the signal source.

The measure has a number of natural properties for a measure of uncertainty:

1. If the message is known, the measure is zero: there is no uncertainty.
2. For a given number of messages, the measure is a maximum when all signals are equally probable. The greatest degree of uncertainty is when no possibility is any more likely than any other.
3. As the number of potential messages increases, this maximum increases. Adding more possibilities increases uncertainty.

It might seem counterintuitive to have an information measure that increases as uncertainty increases. Normally, information might be seen as the opposite of uncertainty. This contradiction is typically resolved by describing uncertainty as quantifying the lack of knowledge about which message is sent while describing information as quantifying the resulting gain in knowledge when the message is received.

The Shannon measure has a further property that uniquely defines it: If a given possibility can be divided into more fine-grained possibilities, the uncertainty over all fine-grained possibilities is equal to the uncertainty of the coarse-grained original possibilities, plus the average of the remaining uncertainty over the additional fine-grained possibilities. Broadly, this corresponds to the idea that it does not matter the order in which fine-grained information is acquired: If you are uncertain which playing card is drawn from a deck of cards, your uncertainty should not depend on whether first you are told which suit it is and then told the number or whether you are simply shown the card straight away.

Acquisition of new knowledge will, on average, reduce uncertainty. However, uncertainty is not necessarily reduced in all cases of acquiring knowledge. The measure does not quantify how correct someone is. Suppose someone is 95% sure that their keys are in their pocket, but equally uncertain of a great many other places the keys might be if not. They are reasonably certain, but not completely certain, where their keys are. Putting their hand into their pocket can reduce the uncertainty to

zero, if the keys are there, but will greatly increase their uncertainty if they do not find their keys.

Information Between Correlated Systems

An important part of information theory is quantifying uncertainty in situations where there is a correlation between two different systems. This involves the joint, conditional, and mutual information.

The joint information about the two systems is simply the uncertainty over all pairs of possibilities for the combined system. The conditional information for one system is the difference between the joint information and the uncertainty for the second system alone. It represents how much uncertainty remains about the first system, once all uncertainty for the second system has been removed. In the case of systems where there are no correlations, the conditional information is equal to the initial uncertainty about the first system. If there are correlations, the conditional information is less than this: Finding out about the state of a correlated system also gives knowledge about the state of the first system.

It should be noted, however that like the uncertainty measure in general, this reduction holds only on average. It may be the case that finding specific data about the second system can leave someone less certain about the state of the first system than they had been beforehand, for example, in the case where the second system is the location of someone's keys and the first system is the place they are going to eat dinner. For perfectly correlated systems, the conditional information is zero: Finding out the state of one system will leave nothing left to know about the state of the other system.

The mutual information is the difference between the sum of the individual uncertainties and the joint information: It quantifies how much information is held in the correlation between the two systems. For uncorrelated systems, it is zero. For perfectly correlated systems, it is equal to the joint information itself.

Operational Treatment of Information

The Shannon information measure, while having properties that seem to capture many intuitive

aspects of information and uncertainty, derives its principal importance from its operational properties. It provides a measure of the ability to efficiently store and transmit information through data compression and error correction.

Data compression refers to the need to encode messages in the smallest number of possible signal states, such as the dashes and dots of Morse code or the ones and zeroes of binary digits. Given a fixed set of possible messages, and a finite number of possible signal states to be used to send the messages, what is the most efficient means of encoding the messages in the signal states?

The noiseless coding theorem shows that the average number of signal states used per message cannot be less than the Shannon information measure of the probabilities of the messages. This theorem applies equally to storing information, such as in the registers on a computer's hard drive, as it does to sending signals over transmission lines. The most efficient schemes for compressing data files in video, audio, or document formats will approach the Shannon measure for how much space they must occupy on a computer.

To achieve the Shannon limit for data compression, the coding scheme must assign a probability to each message, then assign a unique sequence of signal states to that message, in such a way that the more probable messages have the shortest sequences of signal states and the least probable messages have the longest sequences. So, in the English language, the letter E, which is the most frequently occurring, would be represented by a very short code, while the letter Q would tend to be represented by a longer code.

Finding optimal data compression schemes is not straightforward. Although Q is not a frequent occurrence in English, it occurs far more frequently in conjunction with U than in conjunction with X. Truly efficient data compression schemes look for patterns or sequences that are more probable than others. At the extreme, messages might be more efficiently compressed at the level of whole words or even sentences or paragraphs, than individual letters, but practical problems with even all the possibilities, and assigning probabilities to them, rapidly outweigh any gains that might be achieved by such schemes.

The noiseless coding theorem does not take into account the possibility of errors occurring.

Data that are stored or transmitted might be corrupted or distorted by noise. The noisy coding theorem concerns the rate at which messages can be transmitted, without error, along a communication channel which has noise. In a communication channel which has no noise, the signal state transmitted will be the signal state received; there will be a perfect correlation between transmitter and receiver. When there is noise and interference in the channel, a given signal state will have some probability of being corrupted to a different signal state or lost altogether. The correlation between transmitter and receiver will be imperfect. Mutual information measures the correlation between the transmitted and received states.

Error correction codes involve adding extra length to the sequences of signal states to protect the message against this. The simplest example of an error correction code would be one in which a zero is represented by a sequence of three zeroes and a one is represented by a sequence of three ones. If a single one of the transmitted bits is affected by noise and changed to a different value, the remaining two bits still faithfully represent the signal. The receiver can still tell which signal was originally sent by taking the majority. To lose the message, two of the transmitted bits would have to be affected by noise, and this is less likely to happen.

Increasing the length of the signals by using an error correction code will reduce the probability that the wrong message is decoded by the receiver, at the cost of needing more signal states to transmit the same amount of information. The rate of signal transmission measures the average number of signal states transmitted per signal state used in the error correction code. The noisy coding theorem shows that a transmission channel subject to noise has a maximum rate at which signals can be transmitted without error, called the *channel capacity*.

This is given by the mutual information between the signal states and the received states, for an optimal choice of signal probabilities, with respect to the noise in the channel. The information theoretic quantification of the degree of correlation, between transmitted and received signal states, is therefore also the operational quantification of the minimum rate at which the signal states must be lengthened to correct for errors in the messages.

Further Developments

The Shannon information measure is intimately linked to notions of computational complexity. The compressibility of a signal source can be recast in terms of what a computer program can achieve as the encoding and decoding device. Kolmogorov complexity theory relates Shannon information measures to the length of computer programs and the abstract idea of computers as universal Turing machines. In many respects, the theory of computers as Turing machines reflects the same concerns as information theory, separating computer science both from the physical problems of building computers and from the significance of the computations performed.

The operational connection of the information measures to data compression and error correction is clearest when a single signal source transmits down a single line to a single receiver. The quantification of information becomes more complex in more complex scenarios, involving multiple parties, where security, privacy, and authentication are problematical.

The uniqueness of the Shannon information measure depends upon the importance attached to the average cost, in terms of bits, of transmitting messages. However, optimizing a resource cost by an averaging process is not always desirable, particularly in situations where there is a long tail of highly improbable outcomes that may nevertheless have a significant impact on the average. Alternative strategies might involve disregarding very low probability events or might involve adding extra weight to such events to provide protection against catastrophic failures. A family of information measures exists, the Renyi information measures, which include the Shannon information measure as a special case, and generalize the methods of information theory to account for alternative optimization strategies.

The separation of information theory from the physical problems of constructing processing and communication devices has been challenged by quantum theory. Information processing and communication tasks performed by quantum devices can, in some cases, appear to exceed the resource bounds set by information theory. This has led to the development of quantum information theory as a new field of research. Analogues

of the data compression and error correction theorems lead to a quantum measure of information, which contains the Shannon measure as a special case.

Information theory also seems to have a close parallel to statistical theories of thermal physics. The Shannon information measure has the same mathematical form as the Gibbs entropy, and it is commonplace to find physics textbooks referring to thermodynamic entropy as a measure of uncertainty or ignorance.

Owen J. E. Maroney

See also Artificial Intelligence; Church–Turing Thesis; Computer Science; Kinetic (Molecular) Theory

Further Readings

- Aaronson, S. (2013). *Quantum computing since Democritus*. Cambridge University Press.
- Florida, L. (2010). *Information: A very short introduction*. Oxford University Press.
- Nielsen, M., & Chuang, I. (2000). *Quantum computation and quantum information*. Cambridge University Press.
- Shannon, C., & Weaver, W. (1949). *The mathematical theory of communication*. University of Illinois Press.

INFORMATION THEORY, HISTORICAL BACKGROUND

Quantitative information represents objective properties in the natural and social worlds in the form of quantities, numbers, statistics, and frequencies. As opposed to qualitative information, which is concerned with human experience, meaning and motives, quantitative information is the product of mathematical models and hypotheses. Quantitative information begins with empirical observation so as to mathematically express casual and correlative relationships. Assuming that the world is rational and discoverable, quantitative researchers operationalize and isolate concepts by controlling for extraneous influences, then use statistical, mathematical, and computational techniques to analyze numerical data collected from a representative sample to draw lawful

generalizations about a larger population. Following the work of French philosopher Auguste Comte (1798–1857), quantitative information is used to develop testable theories about the products, patterns, and processes of empirical phenomena in the social, behavioral, mathematical, and natural sciences.

This entry first describes the cultural transition from orality to literacy that makes quantitative information possible by abstracting words and numbers from immediate experience and separating logic from the essence of sociality. Second, it discusses Euclidean geometry as a model for grounding the natural and social sciences in quantitative information. Third, it outlines the efforts of Francis Bacon (1561–1626), Galileo Galilei (1564–1642), René Descartes (1596–1650), Isaac Newton (1643–1727), and Comte (1798–1857) not only to legitimate quantitative methods but also to create a modern mentality centered on quantitative information.

Contemporary understandings of quantitative information concern not only the nature and scope of quantitative information and its ability to represent the empirical world but also the capacities and limitations of quantitative methods and their ability to locate and capture information about the social world. Quantitative information combines the notions of measurement and counting, done every day in the real world, as well as the model of functions that grounds Newtonian mechanics and, more recently, statistical data analysis procedures developed in the 19th century.

The dominance of Newtonian mechanics since the 18th century, and at least in colloquial discourse since Einstein in the 20th, promotes a confidence in the explanatory power of quantitative information that grounds the modern world. The spirit of quantitative information presupposes that everything in the empirical world can be measured, and it is the goal and responsibility of modern science to measure it. In the 19th century, the purview of quantitative information was extended beyond necessarily causal phenomena held together by deterministic laws, as in Aristotle's physics, to contingent events associated by probabilistic relationships, traditionally the province of rhetoric.

Together the work of Euclid in the 3rd century B.C.E, Bacon, Galileo, Descartes, and Newton in

the 17th century, and Comte in the 19th century, among many others, have used quantitative information to synthesize ancient and medieval worldviews into a modern world. Descartes's tree of knowledge rooted in metaphysics and Comte's pyramid of knowledge based in mathematics have both legitimated the use of quantitative information as a way to understand the world, and became the backbone of 19th-century science, and the impetus for modern society that revolves around progress, efficiency, and individual autonomy.

From the compatibility with human experience of Euclid's geometry, Bacon's method, Galileo's cosmology, Descartes's doubt, and Newton's mechanics emerged a new modern mentality with quantitative information at the fore. Before explicating the work of Euclid, Bacon, Galileo, Descartes, Newton, and Comte and its influence on the development of quantitative information, let us first describe the cultural conditions that made quantitative information possible in the evolution of words and numbers from primary orality to literacy.

From Orality to Literacy

In *The Presence of the Word* (1967) and *Orality and Literacy* (1982), historian and philosopher Walter J. Ong argues that the consciousness of people with no knowledge of writing differs in fundamental ways from that of literate people and those with knowledge of electronic media. The oral consciousness always lives in the present and commands an immediate audience. The spoken word only makes sense when there is someone present with whom to speak. Oral cultures revolve around narratives, such as Homer's *Iliad* and *Odyssey*, that celebrate and perpetuate the culture's values. In orality, story is history. The storyteller embodies the story and cannot detach himself from it. There are no written records, so culture is embodied as the personal and subjective and performed for others in a repetitive and incremental way. The consciousness of literacy, on the other hand, abstracts words from experience and outsources memory to written records. The writer's audience is always a fiction that emerges from a linear understanding of past, present, and future that privileges objectivity. From this perspective,

the technology of writing alienates the self, exteriorizes thought, and restructures consciousness.

The spoken word in oral cultures differs from the written word in literate cultures in several ways. First, the spoken word is additive rather than subordinative. Whereas the written word obeys strict syntactical and grammatical rules, the spoken word is linear, adding one event to another to aid memory, as in the opening lines of the Book of Genesis in the Douay-Rheims Bible: "In the beginning God created heaven, and earth. And the earth was void and empty, and darkness was upon the face of the deep; and the spirit of God moved over the waters. And God said: Be light made. And light was made."

Second, the spoken word is aggregative rather than analytic. The opening stanzas of the Babylonian epic of Gilgamesh read, "Whose name, from the day of his birth, was called 'Gilgamesh'? Two-thirds of him is god, one-third of him is human . . . beautiful, handsomest of men . . . perfect." These adjectives that pepper speech in oral cultures would be cumbersome and flowery as written sentences. Third, the spoken word is redundant and conservative. Redundancy in speech makes it easier for the rhetor to remember and the audience to follow, as in the African epic of *Sundiata*, which reads, "King of Nyani, King of Nyani, / (Mmm) / Will you ever rise? / King of Nyani, King of Nyani, / (Mmm) / Will you ever rise?" The conservative nature of the spoken word allows oral cultures to pass on their cultural values from generation to generation without written records.

Fourth, the spoken word in orality is close to the human lifeworld and agonistically toned. There are no abstract concepts in orality. The violent imagery of the spoken word in antiquity, as in the climactic scene in Book 22 of the *Odyssey* when Odysseus returns to Ithaca to face the suitors for his wife Penelope, gives meaning to the struggle for survival in the ancient world. Finally, the spoken word is homeostatic and empathetic rather than objectively distanced. Spoken words in oral cultures correspond to one meaning. Without written records, the layers of meaning present in literate cultures are unnecessary.

For Ong, writing is an artificial technology that transforms human consciousness. Writing disconnects words from the existential present. Words, numbers, and their manifestation in the

form of quantitative information become events independent of time and space. Writing separates individuals from the group in space, thought, and work. Letters and numbers are extensions of the sense of touch that fragment and detribalize man. The advent of print permitted replicability and reproduction of abstract units that created a sense of the abstract infinite, which distracts man from his fundamental finitude. The invention of writing changed the social fabric of the West by compartmentalizing time into measurable fractions. Events in time no longer created their own time and space but were contained within time and space. Preliterate cultures experience time; they do not measure it.

Also, existential settings infuse words with meaning in ways that dictionaries cannot. In oral cultures, auditory utterance shapes and determines the cognitive thought processes of the speaker and audience. All speech is shaped by an anticipated response. The technology of writing separates words from their author by separating logic from rhetoric, the knower from the known, and sight from sound. The written word is voiceless and cannot be directly challenged or contested like oral speech. This detachment occurs in oral cultures when the speaker is considered only the channel, not the source of oral speech, as in oracles or religious rituals. Critiques of the effects of technology on culture, most notably by Plato in the *Republic*, were themselves made possible by the effects of writing on mental processes.

Ong argues that writing is unresponsive, artificial, and manufactured. Writing weakens the mind and destroys memory. The distance between a text and its author sharpens the analytical capacity of individual words and numbers. Printing spawns the need for listing and indexing that further detaches and abstracts the words of a text from their situational and linguistic context. The tangibility and visual nature of a printed text, as opposed to the fleeting and existentially present spoken word, promotes knowledge as an illustration in space, independent of time and source. According to Ong, the printed word allows for an objectification of human consciousness that, in turn, fosters interiority and privacy that did not exist in action-oriented oral cultures.

The interaction of writing and orality, and the impact of residual orality, is manifested clearly in

the development of rhetoric, the art of oral address for persuasion. Because rhetoric cannot be detached from the human lifeworld, rhetoric characterizes the entire human experience, not just public speech. Rhetoric is not only a means of persuasion but also the means of identity affirmation. The emphasis on rhetoric as the essence of sociality demonstrates its close connection to the human lifeworld. For Ong, quantitative information has no meaning outside of a shared social context in the rhetorical arena of lived experience.

Euclidean Geometry

The evolution of the word from orality to literacy in the ancient Greek world made possible the evolution of geometry from a practical and economical means of measurement, to a deductive science that sought knowledge for its own sake. Euclid's axiomatic-deductive geometry is representative of this trend in the 3rd century B.C.E. and continues to influence the infrastructural foundations of quantitative information in both the natural and social sciences.

Euclid outlines his approach to geometry in the *Elements*, written in Alexandria around 300 B.C.E. In Euclidean geometry, axioms are presented as self-evident descriptors of natural reality. From those axioms, deductive logical inferences are used to prove theorems using only a ruler and a compass. Notwithstanding the development of non-Euclidean geometry and Einstein's relativity developed in the 19th and 20th centuries, Euclid's *Elements* has formed the foundation for quantitative information in the natural and social sciences. The *Elements* contains a series of axiomatic postulates, from which theoremetric propositions are deduced through mathematical proofs.

The contents of the book range from the postulate of plane geometry, including geometric algebra and the properties of circles, triangles, and polygons, to ratios and proportions, including magnitudes, divisibility, series, and sequences, to spatial geometry, including pyramids, perpendicularity, and parallelism. In the *Elements*, Euclid first defines a point—that which has no part—and a line—a length without breadth—then defines an angle as the inclination of two straight lines, and a circle as a plane figure consisting of all points that have a fixed distance from a given center.

These definitions in Euclidean geometry apply to flat space, where the distance between two points in two or three dimensions can be calculated by an axiomatic-deductive formula. For instance, when drawing on a flat piece of paper, it is easy to observe that the shortest distance between two points is one unique straight line, that the sum of the angles in any triangle is 180° , or that through a point not on a line there is no more than one line parallel to the line.

The oldest extant treatise of axiomatic-deductive mathematics, the *Elements* was among the first texts to be typeset after the invention of the printing press, second only to the Gutenberg Bible in number of editions printed. Until the second half of the 19th century, geometry was synonymous with the Euclidean geometry taught in secondary schools. Euclidean geometry allowed one to apply theorems in diverse situations by generalizing centrifugally from a few known facts. Although other non-Euclidean geometries developed in the 19th and 20th centuries have added to Euclid's understanding of finite space, Euclidean geometry remains the most compatible conception of space with how human beings experience the world. This emphasis on proof through insight rather than memorization infused quantitative information with not only descriptive but also generative explanatory power.

The Scientific Revolution as the Spirit of Quantitative Information

With the publication of *On the Revolutions of the Heavenly Spheres* in 1543, Polish astronomer Nicolas Copernicus sparked a scientific revolution that would culminate with Newton's *Mathematical Principles of Natural Philosophy* in 1687. In his book, Copernicus laid out his heliocentric vision of the universe with the sun, not the earth, at the center of the solar system. This view of the universe was directly at odds with the Ptolemaic cosmology based on Aristotelian mechanics, with the earth at the center of the solar system surrounded by the other three basic elements—air, water, and fire—revolving in concentric circles. Copernicus's heliocentric cosmology took the place of Ptolemy's geometric cosmology but provided no replacement for Aristotle's mechanics. Without this replacement, the Copernican system

was counterintuitive. If the earth is spinning on its axis and revolving around the sun, why does it appear at rest to the casual observer?

Italian philosopher, astronomer, and mathematician Galileo Galilei induced an answer from his experiments with balls rolling on incline planes. In his *Discourses and Mathematical Demonstrations Relating to Two New Sciences* (1638), Galileo argued that Aristotle was wrong to believe that unnatural motion required a proximate cause. Instead, an object in motion, like the earth, will tend to stay in motion through inertia. Because everything on earth is moving contemporaneously, the movement is imperceptible to the casual observer. After French philosopher and mathematician Rene Descartes extended Galileo's principle to motion in straight lines, Isaac Newton restated it as his first law of motion in the *Principia*.

Galileo's observations using the telescope led him to accept in theory the Copernican system. However, Copernicus still adhered to some Platonic ideals that left many unanswered questions in his heliocentric vision. German mathematician and astronomer Johannes Kepler (1571–1630) devoted his career to answering these questions. Kepler proposed that the orbits of the planets were not circles, as Plato argued, but ellipses around the sun that could be calculated as a function of the planet's distance from the sun. Newton's universal law of gravitation and postulates on force, mass, and motion were a synthesis of the work of Copernicus, Descartes, Galileo, and Kepler, and at last provided the replacement for Aristotle's mechanics within Copernican cosmology.

Galileo's Defense of Quantitative Information

The work of Galileo legitimated the use of observation, measurement, and experiment now synonymous with scientific method and quantitative information. In *The Assayer* (1623), Galileo argued that the language of mathematics can explain the growth of the smallest plant and the workings of the entire cosmos with the same precision. His words in *The Assayer* are worth quoting at length: "Philosophy [i.e., physics] is written in this grand book—I mean the universe—which stands continually open to our gaze, but it cannot be understood unless one first learns to

comprehend the language and interpret the characters in which it is written. It is written in the language of mathematics, and its characters are triangles, circles, and other geometrical figures, without which it is humanly impossible to understand a single word of it; without these, one is wandering around in a dark labyrinth."

In 1609 in Padua, Galileo began observing the heavens with his 20-powered spyglass telescope. He observed that the surface of the moon was rough and cratered, not smooth and even, that four moons revolve around Jupiter, and that many more stars exist than are visible by the naked eye on Earth. Galileo later discovered the anomalous appearance of the planet Saturn, and the phases of the planet Venus. The discovery of craters on the moon contradicted Aristotle's cosmology, in which the imperfect earth existed below the perfect Heavens. The discovery of the moons orbiting Jupiter showed that the earth could not be the only center of motion in the universe. The discovery of the phases of Venus showed that Venus revolves around the Sun and led to Galileo's conversion from Aristotelian to Copernican cosmology, in which the sun, not the earth, is the center of the universe.

Aristotelian objectors to Galileo argued that science concerns universal laws, not particular observations, and that a science of particulars grounded in observation, measurement, and experiment can achieve no knowledge of the universal and unobservable laws of physics. Galileo responded to Aristotelians who argued that measurement and observation tell us nothing about the essence of natural phenomena by arguing that God set the world in motion through geometrical laws, and that these laws are knowable through observational measurement. For instance, the imperfections on the surface of the moon visible through Galileo's telescope are incompatible with the Aristotelian law that all planetary bodies are perfectly spherical.

Galileo's elaboration on Aristotle's idea of induction also has implications for quantitative information. In the *Posterior Analytics* (350 B.C.E.), Aristotle described induction as a cognitive process whereby one argues from a particular to a universal to arrive at the first principles of science. Galileo developed Aristotle's idea of induction as a justification for grounding scientific

knowledge in observation. One proves the truth of the universal laws that ground empirical sciences through inferences made from many observations. Galileo's defense of quantitative information from Aristotelian ontologists legitimated the use of quantitative methods in natural philosophy and astronomy.

Quantitative Information and the Modern World

Francis Bacon

Francis Bacon's faith in the perfection of the scientific method and experience as the sole source of knowledge has guided the spirit of quantitative information. Bacon's *New Organon* (1620) was a comprehensive re-legitimation of natural science using natural philosophy to reform scientific method. Bacon sought to create a new method of inducing universal laws from particular observations. Bacon's "tables of discovery" provided new ways to process and make sense of particular human experiences. By categorizing particular experiences using these tables, Bacon argued that one could induce physical causes and metaphysical essences of phenomena.

In the *New Organon*, Bacon outlined his method for studying facts to interpret natural phenomena as an alternative to Aristotle's logic in the six books of the *Organon*, which subordinated natural philosophy to syllogistic logic. As outlined in the *New Organon*, Bacon's method consisted of four steps. First, one brackets all presuppositions and preconceived notions that could influence the objectivity of interpretation. Next, one describes the facts. Third, one uses the tables of discovery to categorize the facts according to the presence, absence, and degree of the characteristic under study. In this step, one makes a list of things or situations in which the phenomenon occurs and does not occur and the degree to which it occurs in each list. Once the facts are categorized, one eliminates all the characteristics that are associated with the phenomenon in order to isolate the cause of the phenomenon. From interpreting these lists, one can deduce the form, or underlying cause, of the phenomenon. This hypothesis becomes one of many such hypotheses that by

gradual degrees get closer to the true cause of a phenomenon.

Bacon's method and philosophy formed the intellectual foundations of modernity. His unflinching faith in the explanatory power of quantitative information remains ingrained in the masses, and his understanding of experience as the sole source of knowledge remains ingrained in the academy. Bacon's rejection of Aristotle's idea of formal cause, where change is caused by the form of the thing that changes, paved the way for modernity's repudiation of tradition, whereby the accumulation of facts through experience is what leads to human knowledge acquisition.

If Bacon is the father of modernity, then the guiding narrative of modernity is not the *polis*, as in classical antiquity, or the church, as in the Middle Ages, but faith in progress, efficiency, and individual autonomy. The Baconian method embodies the contradictions of the modern world; that is, that mastery over the natural world is a Faustian endeavor. Bacon transforms reason from a defining quality of the human condition in Aristotle, to a means to achieve the end of scientific progress in the form of quantitative information. Bacon sought to liberate humanity from the quicksand of tradition and the ancient and medieval "cult of books" to create the conditions for a modern world that revolves around objective experimentation. Bacon's emphasis on the utility of knowledge and power over nature led to a new quantitative economy that bridges the gap between an ancient and medieval economy that is naturalistic, organic, qualitative, and concrete to a modern economy that is materialistic, mechanistic, quantitative, and abstract.

René Descartes

Together with Bacon, René Descartes has influenced the intellectual trajectory of quantitative information. In his *Discourse on the Method* (1637), Descartes attempts to arrive at a fundamental set of principles that one can know as true without any doubt. Through systematic and methodological doubting laid out in the *Discourse*, one of the first philosophical texts written in the vernacular French and not in Latin, Descartes concludes in the *Meditations on First Philosophy* (1641) that if he doubted, then something or

someone must be doing the doubting, and therefore the very fact that he doubted proved his existence. Descartes argued he can be certain that he exists because he thinks: *cogito, ergo sum*. Descartes held that the mind alone was capable of grasping the truth with mathematical certainty, and morality was a science with its roots in metaphysics.

For Descartes, experiment, scientific inquiry, and action trump disputation, rhetoric, and speculation. These emphases became the backbone of 19th-century science and the impetus for modern society. Rhetoric itself could not produce truth with mathematical certainty and was relegated to communicating the principles that only science could discover. Rhetoric disturbs the clarity of rational thought and is nonphilosophical in its preoccupation with contingencies rather than eternal truths.

Upon the foundations of Euclidean geometry, Descartes synthesized algebra and geometry to develop analytic geometry. Using Descartes's analytic *geometry*, as articulated in *Geometry*, an essay accompanying the *Discourse*, one can formulate equivalent problems using algebraic equations and geometric shapes by manipulating their mathematical languages. Descartes's coordinate system allows one to map the locations of geometric forms such as points and lines using pairs of quantitative coordinates. The distances, areas, and volumes of geometry could be defined using numbers and expressed using equations.

Descartes aimed to re-legitimize philosophical discourse and scientific method in response to skeptics such as the Greek physician and philosopher Sextus Empiricus and, later, French essayist Michel de Montaigne, who argued in his *Essays* (1580) that the sources of human knowledge are unreliable. In the *Meditations*, Descartes reflects on all of the falsehoods he has believed in his life and all of the beliefs that could be proven false; he aims to start from scratch with what he can be absolutely certain of with no doubt. He argues that any belief that is uncertain or potentially proven wrong must be discarded. For instance, authorities and tradition cannot be trusted because authorities are frequently wrong. The senses cannot be trusted because sensible objects often appear skewed under water or from a distance. Immediate experience cannot be trusted because it may be a dream. Finally, mathematical truths such as $2 + 2 = 4$ cannot be trusted

because God may have created imperfect conceptions of mathematics. Even if there is no God, it is more likely that the senses would be imperfect if they were not created by a perfect Being.

Descartes concluded that the only thing he could be absolutely certain of with no doubt was the intuition that, even if he was being deceived, that someone or something was thinking. For Descartes, then, the *cogito* is the only self-evident truth. However, any other idea that proves the existence of something as clearly and distinctly as the *cogito* proves the existence of the self cannot be doubted. In his epistemological project in the second part of the *Discourse*, Descartes separates the immortal mind and the mortal body and argued that anyone can use reason to discover the truth. Knowledge is like a tree with its roots in metaphysics, trunk in physics, and the branches as the higher forms of scientific knowledge that house the fruit of knowledge. In the third part of the *Discourse*, Descartes offers a provisional—and later final—moral code that requires that one obey the laws and the customs of one's country and religion, be firm and resolute in holding true to the opinions one adopts, change oneself before attempting to change the world, and always chase the truth using the method of universal doubt.

Isaac Newton

Isaac Newton adhered to Descartes's code. As a cosmological framework, Newtonian mechanics undergirds the belief in scientific progress as the guiding narrative of modernity. Newtonian mechanics explains the motion of bodies according to the gravitational, electrical, and magnetic forces exerted on bodies when in motion and at rest. Just as the staying power of Euclidean geometry into the 19th century was due in part to its consistency with human experience, the success of Newtonian mechanics in accurately describing the motions of phenomena from celestial bodies to ordinary objects on earth and to atomic and subatomic matter is due in part to its compatibility with human intuition.

In his *Principia*, Newton proposed his laws of motion and equilibrium, which would become the laws of conservation of energy, momentum, and angular momentum central to modern physics.

Newton's first law is the principle of inertia proposed by Galileo and tweaked by Descartes. Newton's second law concerns force, mass, and acceleration and produced one of the most famous equations in history, F (Force) = m (mass) a (acceleration). Newton's third law states that for every action on a body at rest, in uniform motion or in accelerated motion, there is an equal and opposite reaction. Together, these laws set forth in the *Principia* form the basis of Newtonian mechanics, which, as a system, has provided the language and the intuition of scientists for three centuries.

Comte and Social Science

Guided by the spirit of the scientific revolution spearheaded by Galileo, Bacon, Descartes, and Newton, French philosopher Auguste Comte grounded not only the natural world but also the social world in quantitative information. Comte draws parallels between the intellectual maturations from infancy to adulthood of individuals and all of humankind as a model for the sciences.

For Comte, this maturation advances in theological/mythical, metaphysical/philosophical, and scientific/positive stages. In the theological/mythical stage, supernatural deities explain natural phenomena. In this stage, humans attribute the causes of unexplainable phenomena to anthropomorphic and divine beings in the image of themselves. The theological/mythical stages consist of three sub-stages: from fetishism, polytheism, and monotheism, representing the development of knowledge from worshipping inanimate objects to worshipping many gods of different natural phenomena to worshipping one omniscient God. In the metaphysical/philosophical stage, abstract and impersonal concepts define natural phenomena. For instance, Plato contrasts the imperfect world of appearances with the perfect realm of ideal Forms. The metaphysical/philosophical stage abstracts God from the world and replaces Him with an abstract God that directs the actions of the world.

In the scientific/positive stage, empirical observations, informed hypotheses, and replicable experiments explain natural and social phenomena. In this stage, humans induce laws from patterns and regularities in nature and

society and use the descriptive laws to make sense of particular cases. For Comte, humankind only reaches intellectual maturity in the third stage by discarding theological explanations for phenomena in the natural and social worlds and committing to the logic and rationality of the scientific method. Whereas Descartes viewed knowledge as a tree with metaphysical roots and a physical trunk supporting the branches of science, Comte arranged the sciences into a pyramid with mathematics at the base, followed by the physical, natural, and cerebral sciences, and culminating in social science. Comte saw himself as rescuing *sociology* from the uncertainties of philosophy and linking it to the certainties of science.

David DeIuliis

See also Information Theory; Paradigm; Rationality; Reasoning

Further Readings

- Biener, E., & Schliesser, E. (2014). *Newton and empiricism*. Oxford University Press.
- Bonet, E., Jensen, H. S., Iglesias, O., & Sauguet, A. (2007). *Background of the debate on qualitative and quantitative methods* [Paper presentation]. Seventh Annual Conference of the European Academy of Management, Paris, France. https://academia.edu/176086/Background_of_the-debate_on_qualitative_and_quantitative_methods
- Dear, P. (2009). *Revolutionizing the sciences: European knowledge and its ambitions, 1500–1700*. Princeton University Press.
- Hanson, B. (2008). Whither qualitative/quantitative? Grounds for methodological convergence. *Quality and Quantity*, 42(1), 97–111.
- Jardine, L. (1999). *Ingenious pursuits: Building the scientific revolution*. Random House.
- Kuhn, T. (1962). *The structure of scientific revolutions*. University of Chicago Press.
- Mackenzie, N., & Knipe, S. (2003). Research dilemmas: Paradigms, methods and methodology. *Issues in Educational Research*, 16(2), 195–205.
- Ong, W. J. (1982). *Orality and literacy: The technologizing of the word*. Methuen.
- Rosenberg, D. (2003). Early modern information overload. *Journal of the History of Ideas*, 64(1), 1–9. <https://doi.org/10.1353/jhi.2003.0017>

- Shah, S. K., & Corley, K. G. (2006). Building better theory by bridging the quantitative–qualitative divide. *Journal of Management Studies*, 43(8), 1821–1835. <https://doi.org/10.1111/j.1467-6486.2006.00662.x>
- Vickers, B. (1992). Francis Bacon and the progress of knowledge. *Journal of the History of Ideas*, 53(3), 495–518. <https://doi.org/10.2307/2709891>

INSTRUMENTALISM

According to instrumentalism, the value of any particular theory does not rest in its ability to capture truth but in its usefulness for understanding and predicting phenomena. On this view, the real worth of a theory is not in its truth or approximation of reality but in its value as a useful instrument in solving practical problems. The natural and social sciences are widely appreciated because of their capacity to provide solutions to many of the problems humans face. Even for those issues humans are not sure science has been able to explain, there is the belief that science will one day achieve breakthroughs and those problems will be solved. This entry traces the origins of instrumentalism and its engagement with pragmatism and scientific realism and concludes with a brief account of the emergence of a related development in scientific theory, namely, constructive empiricism.

Origin and Significance

Instrumentalism and the related tradition of pragmatism are associated with the American philosophers William James (1842–1910) and John Dewey (1859–1952), whose works greatly influenced philosophy, psychology, education, and civil society. Pragmatism understands that knowing the world is inseparable from agency within it, that philosophical concepts should be tested by scientific experimentation, and that a claim is true only if it is useful. According to Dewey, instrumentalism is the position that all the acceptance of a given theory amounts to is a judgment about the theory's usefulness in predicting phenomena. Relatedly, scientific realism holds that the world as described by science is the actual real world, regardless and independent of what humans imagine and think it is.

Although instrumentalism is in constant engagement with both pragmatism and scientific realism, it is different from them as it considers, largely, that ideas are appreciated by their transformative capabilities as instruments of change with positive consequences. Instrumentalism is methodological and outcomes focused. It is about doing and action rather than abstract thoughts. According to instrumentalism, theories are not actually truth-evaluable, and they do not always correctly depict reality. The truth of an idea is thus determined by its success in actively providing solutions to problems, while its real value is determined by the roles of ideas in enhancing human experiences.

Scientific Realism and Instrumentalism

A closer consideration of scientific realism may be helpful for understanding instrumentalism. *Realism* holds that there are mind-independent objects. *Scientific realism* is the idea that scientific theories say definite things about genuinely existing entities and that scientific theories get successively better and better at giving a true account of the world. Hence, its reliance on abductive reasoning, or “inference to the best explanation,” has led to the success of science in providing true, or approximately true, descriptions of the world. In contrast, instrumentalism maintains that even a successful scientific theory discloses nothing that is either true or false about unobservable processes or objects; rather, a theory is only a tool for predicting observations, which lead in turn to the formulation of “laws” that these phenomena supposedly obey. For instrumentalists, there is a world of difference between observable and unobservable entities even though they concur with the realists' explanation of observable entities. The major difference lies in the unobservable entities, because instrumentalism considers unobservable entities as merely posited entities based on theoretical considerations. For scientific realists, it is impossible to draw a distinction between the two, for they lie on a continuum and are mutually reinforcing in any effort to make sense of the world. Furthermore, for scientific realists, even when entities and ideas do not serve an immediate practical purpose, this does not diminish their philosophical relevance. Indeed, to reduce the ideas and entities to the level of *tools* diminishes the philosophical

relevance of ideas and negates the role of the philosopher in making sense of the world.

Among the better-known scientific realists is the physicist Wolfgang Pauli (1900–1958), famous for his *exclusion principle*. In essence, according to this principle, no two fermions (elementary particles with “half-integer spin”) in an atom can have the same four quantum numbers; in other words, only one fermion can occupy a particular quantum state at a given time. After decades of systematic observations, many scientists have come to accept that the Pauli exclusion principle is true, although other scientists have not.

Within the philosophy of science, instrumentalism has benefited from the thinking of late 19th- and early 20th-century thinkers such as the physicists Ernst Mach (1838–1916), who believed that physics should be based entirely on directly observable phenomena, and Pierre Duhem (1861–1916). Duhem, in his book *The Aim and Structure of Physical Theory* (1906), held that the physical sciences should be practiced independently of deep metaphysical assumptions. That is, the aim of physical theory is to propose mathematical laws that predict experimental outcomes as precisely and simply as possible. It’s not necessary to believe that the fundamental physical laws are true but rather that they are merely useful fictions that help scientists predict phenomena.

Later Developments

Instrumentalism is dynamic and has continued to develop over time. Even in the face of philosophical antagonisms, what has persisted is the central thesis of instrumentalism that sees theories as tools for pursuing practically useful purposes rather than as correct and true descriptions of the world and nature. On those grounds, instrumentalism is an extreme view of how to think about the nature of scientific theories, as evaluable only in terms of their use as tools of inference. More generally, anti-realism doesn’t commit itself to the existence or reality of the things mentioned in theoretical descriptions but values the theories for other reasons having to do with empirical adequacy, utility, and any number of other theoretical virtues not involving a belief in the entities of the theory.

Constructive empiricism is a prominent version of this class of interpretation of theories. Developed and discussed by the philosopher of science Bas van Fraassen, it became widely known after the publication of his book *The Scientific Image* in 1980. The main contribution of van Fraassen through constructive empiricism is that of actually finding a balance between scientific realism and instrumentalism. In his position on theories as tools, he insisted that just as scientific realism is right in its metaphysical view of scientific theories, so instrumentalism is also correct and equally reasonable epistemically.

According to van Fraassen, the goal and aim of science is to provide theories that are empirically adequate. For a theory to be empirically adequate, it must be right about the observable phenomena. Indeed, many of the theories that are adopted by scientists are only adopted for their empirical adequacy in serving as explanatory frameworks and not necessarily because one believes in them entirely as true. Furthermore, constructive empiricism holds that scientific theories are semantically literal (i.e., its propositions are either true or false) and are adequate only if everything they say about observable entities is true.

Perhaps the chief difference between early and contemporary instrumentalism is a matter of emphasis. For instance, while Duhem’s strand of instrumentalism insisted that the objective and purpose of physical theories ought to be instrumental, van Fraassen holds that the aim and objective of physical theory could actually be instrumental. van Fraassen’s contemporary instrumentalism has thus only balanced the philosophical axis and continuum of both scientific realism and instrumentalism, particularly through the position that it is also about the rationality of the philosophers in question and their capacities for philosophical accommodation. Thus, van Fraassen preserved the original view of instrumentalism, maintaining the view that instrumentalism remains largely the best approach to making sense and understanding science philosophically.

Olayinka Akanle

See also Epistemology; Explanation; Theory Construction

Further Readings

- Dewey, J. (1916). *Essays in experimental logic*. The University of Chicago Press.
- Duhem, P. (1954). *The aim and structure of physical theory* (W. Philip, Ed.). Princeton University Press.
- Hendry, R. F. (2001). Are realism and instrumentalism methodologically indifferent? *Philosophy of Science*, 68(3), S25–S37.
- Kukla, A. (1992). On the coherence of instrumentalism. *Philosophy of Science*, 59(3), 492–497.
- McCullagh, C. B. (1973). Historical instrumentalism. *History and Theory*, 12(3), 290–306.
- McGinn, M. (2013). Instrumentalism and poetic thinking: A critique of Dewey's logic of thought. *Stance*, 6, 45–52.
- Otto, M. C. (1926). Instrumentalism. *The Monist*, 36(4), 577–593.
- Papineau, D. (1986). The paradox of instrumentalism. *PSA: Proceedings of the Biennial of the Philosophy of Science Association*, 1, 269–276.
- Psillos, S. (1999). *Scientific realism: How science tracks truth*. Routledge.
- Schiller, F.C.S. (1925). Instrumentalism and idealism. *Mind*, 34(1), 75–79.
- van Fraassen, B. C. (1980). *The scientific image*. Clarendon Press.

INTERPRETATION

The *interpretation* of a theory assigns meaning to some or all components of the theory by connecting them to independently interpreted statements or theories, or directly to the world. As this entry explains, the precise kind of connection depends on whether the theory is mathematical or empirical, on whether it is expressed informally or formally, and on the kind of formalism. The interpretation of a mathematical theory connects it to another theory, typically with the goal of showing how specific formal properties from one of the theories transfer to the other. Such an interpretation is well defined and uncontroversial. The interpretation of an empirical theory, on the other hand, is an empirical phenomenon that relates not only to semantic aspects like the content of the theory and its conditions for confirmation but also to pragmatic aspects

like the behavior of the scientists and the teaching and learning of the theory. Accordingly, an account of the interpretation of empirical theories needs to relate to these semantic and pragmatic aspects, and there is significant debate about every extant account of the interpretation of empirical theories.

The Interpretation of Mathematical Theories

A mathematical theory is typically interpreted by relating its expressions to other mathematical expressions. Such an interpretation may proceed by formal semantics or by the interpretation of one theory in another. Therefore, this kind of interpretation stays within the purely formal realm. *Interpretation* sometimes also refers to the connection of a mathematical theory to genuine mathematical objects, which are not purely formal but also not part of the physical world.

Formal Semantics

If a theory is formally axiomatized in predicate logic, it can be easily interpreted by interpreting its constant, predicate, and function symbols—in other words, its *vocabulary*—in the formal semantics developed by Alfred Tarski. A very simple theory in predicate logic may contain only the single sentence $Pa \wedge \exists x Qxa$ (“ a is P and there is an x such that x has the relation Q to a ”), where a is a constant symbol, P is a one-place predicate symbol, Q is a two-place predicate symbol, and x is an existentially quantified variable. Thus, the theory's vocabulary consists of a , P , and Q . A structure for the theory in formal semantics then consists of a domain and an interpretation. The domain is a set of objects like $\{\alpha, \beta, \gamma\}$. The interpretation assigns to each constant symbol an element of the domain; a could, for instance, be assigned α . To each one-place predicate symbol, the interpretation assigns a subset of the domain; P could, for instance, be assigned $\{\alpha, \beta\}$. To each two-place predicate symbol, the interpretation assigns a set of pairs from the domain; Q could, for instance, be assigned $\{\langle \alpha, \beta \rangle, \langle \beta, \beta \rangle, \langle \gamma, \alpha \rangle\}$. The structure then determines whether the theory is true or false, which is the theory's *truth value*.

The theory is true if and only if all its sentences (in this case just the one sentence) are true. The truth value of each sentence is determined by the truth values of its subsentences, in this case Pa and $\exists x Qxa$, and the logical connectives connecting them, in this case “ $\wedge$.” Pa is true if and only if the object assigned to a is an element of the set assigned to P , which it is, since $\alpha \in \{\alpha, \beta\}$. $\exists x Qxa$ is true if and only if there is an object in the domain such that the pair consisting of that object and the object assigned to a is an element of the set assigned to Q , which it is, since $\langle \gamma, \alpha \rangle \in \{\langle \alpha, \beta \rangle, \langle \beta, \beta \rangle, \langle \gamma, \alpha \rangle\}$. The whole sentence $Pa \wedge \exists x Qxa$ is a conjunction, and a conjunction is true if and only if both its conjuncts are true, which in this case they are. Thus, the sentence is true in the structure, and the theory consisting of only this sentence is also true in this structure, which is therefore called a model of the theory. (For this reason, formal semantics is also called model theory.) More complicated theories can also involve function symbols and higher order predicate symbols, that is, predicate symbols that apply to constant, predicate, and function symbols rather than only to constant symbols. The basic idea of formal semantics stays the same for such more complicated theories.

If a mathematical theory is well established, like the theory of the natural numbers, for instance, then it is often taken to have a standard model. In the case of the natural numbers, the domain of this standard model is just the set N of natural numbers, the constant symbol “42” is interpreted by the number 42, the two-place predicate “ $>$ ” is interpreted by the set of pairs in which the first is greater than the second, the function symbol “ $+$ ” is interpreted by the function that maps each pair of natural numbers to the sum of its components, and so on. The domain and the entities that interpret the vocabulary of such a standard model are then sometimes taken to exist as genuine mathematical objects, independently of any formalization. They may be taken to exist, for instance, as platonic objects or as objectively real and abstract products of the human mind, as was suggested by Karl Popper. A theory of, say, geometry may thus be interpreted by points and lines in a platonic space or Popper’s world of objective ideas. If the points and lines were part of the physical world, however, the theory would be empirical.

The Interpretation of Theories in Other Theories

A mathematical theory can also be interpreted more generally by mapping its expressions to the expressions of another mathematical theory and by mapping its structures to the structures of the other mathematical theory. Such an interpretation is often used to infer that one of the theories has a specific feature, such as consistency, if the other one does. A theory of geometry, for instance, can be interpreted by vectors of real numbers. Another standard example is the interpretation of the complex numbers as pairs of real numbers. Here, pairs of real numbers are mapped to complex numbers, and every expression referring to a complex number is mapped to an expression referring to a pair of real numbers. The expression about the complex numbers is then true if and only if the expression for the real numbers is true. Since two real numbers make up one complex number in this interpretation, the interpretation is called *two-dimensional*.

More generally, n -tuples (ordered sets with n elements) of one Class A of structures can be used to interpret the elements of the domains of another Class B of structures. This interpretation then leads to an n -dimensional interpretation of the theory of Class B in the theory of Class A .

The Interpretation of Empirical Theories

An interpretation of an empirical theory consists of a connection between the theory and the empirical world; specifically, an interpretation describes what the world would be like if the theory were true. The interpretation of a theory expressed informally in natural language is typically simply taken over from the interpretation of the natural language, but the theory may also be interpreted in analogy to the interpretation of a formally expressed theory. The interpretation of a formally expressed theory can conversely consist of the translation into a natural language, whose interpretation thereby also applies to the theory. (This is an informal analogue to the interpretation of one mathematical theory in another; instead of mapping the expressions of a formal theory to expressions of another formal theory, the expressions of a formal theory are mapped to expressions of natural language.) But a formally expressed theory can also be

interpreted formally, in which case the kind of interpretation depends on the formalism used.

Syntactic Theories

A theory syntactically described in predicate logic can be formally interpreted in model theory. But since an empirical theory has to be connected to the world in some way, the structures interpreting the theory either have to be in turn connected to the world through some interpretation, or the structures themselves must be what may be called *worldly*. For a *worldly structure*, at least some elements in its domain must be objects of the world or some of its sets of objects or sets of tuples of objects must correspond to properties or relations that can be found in the world. In the simplest case, the domain contains only objects of the world, each predicate symbol is interpreted by a property of or relation between the objects in the world, and each function symbol is interpreted by a function that holds between objects of the world. The theory is then true if and only if these objects, properties, relations, and functions make up a worldly structure in which the theory is true.

Already this simplest case gives rise to a central problem for the interpretation of theories, the impossibility of a theory fixing its own *intended interpretation*. The reason for this impossibility is known as Putnam's *model-theoretic argument*, which relies on the isomorphism theorem. Roughly, two structures are *isomorphic* when one can be obtained from the other simply by uniquely exchanging the objects in their domain and in the set-theoretic entities the interpretation assigns to the vocabulary. Thus, the structure mentioned earlier with domain $\{\alpha, \beta, \gamma\}$, which interprets a by α , P by $\{\alpha, \beta\}$, and Q by $\{\langle \alpha, \beta \rangle, \langle \beta, \beta \rangle, \langle \gamma, \alpha \rangle\}$, is isomorphic to the structure with domain $\{\beta, 2, k\}$ that interprets a by β , P by $\{\beta, 2\}$, and Q by $\{\langle \beta, 2 \rangle, \langle 2, 2 \rangle, \langle k, \beta \rangle\}$. This is because the second structure can be obtained from the first by replacing α by β , β by 2 , and γ by k . This mapping between the two sets $\{\alpha, \beta, \gamma\}$ and $\{\beta, 2, k\}$ is a bijection (also called a *one-to-one and onto mapping*). Whenever two sets have the same number of elements and that number is greater than 1, there are multiple bijections between these sets. In particular, any element of the one set can be mapped to any element of the other set in some bijection.

And as the earlier example indicates, for every bijection between the domain of a Structure A and some Set S , there is another structure that has the domain of S and is isomorphic to A . For instance, a might not only be interpreted by β but also by 2 or by k , depending on the bijection between the two domains. The isomorphism theorem now states that if a theory is true in some Structure A (has the Model A), then it is true in every structure that is isomorphic to A . And this entails that every true theory has infinitely many models, and in particular every constant, predicate, and function symbol could be interpreted by something completely different: Even if a is intended to be interpreted by α , there is another model in which it is interpreted by β , another in which it is interpreted by 2 , another in which it is interpreted by k , and yet another in which it is interpreted by Liechtenstein.

Because of the model-theoretic argument, the intended interpretation of a theory needs to be fixed at least to some extent without recourse to sentences of predicate logic, for instance, by simply pointing to the objects by which a symbol is to be interpreted. This is known as an *ostensive interpretation* (or sometimes, confusingly, an *ostensive definition*). Ostensive interpretations are obviously only feasible for symbols that are intended to be interpreted by objects that can be pointed at. For predicates which apply to sets that are so large that it is not feasible to point at all their elements (say, infinite sets), ostensive interpretations have to be complemented. This complementation often takes the form of psychological similarities between those who teach, say, a predicate and those who learn it so that the grouping of objects the teacher intends to convey is picked up by the student. However, such complementations typically lead to areas of vagueness so that for some objects it is not clear whether they do or do not fall under the predicate. Fortunately, the formalism of standard model theory can be generalized to allow for vague interpretations. In a standard interpretation, a one-place predicate symbol is assigned a single set of objects, all of which have the property corresponding to the predicate symbol; a many-place predicate symbol is assigned a set of tuples of objects, all of which stand in the relation corresponding to the predicate symbol. In a vague interpretation, each predicate symbol is assigned two

disjoint sets: One containing (tuples of) objects that clearly have the corresponding property (stand in the corresponding relation), and one containing (tuples of) objects that clearly do not have the property (do not stand in the corresponding relation). For (tuples of) objects in neither of these two sets, it is vague whether they have the property (stand in the relation). Similar generalizations exist for function and constant symbols.

In the *received view of theories*, which assumes that the theories are syntactically expressed, the interpretation of a theory is itself described to a large extent syntactically: The logical relations of the theory's vocabulary to another independently interpreted vocabulary is described in predicate logic. The predicate-logical sentences describing these relations are often called *correspondence rules*, but are also known as *coordinating definitions* (although they are typically not definitions). The interpretation of the theory's vocabulary (the *theoretical vocabulary*) is then determined indirectly by the interpretation of the independently interpreted vocabulary and the correspondence rules. If the correspondence rules provide explicit definitions for every predicate, function, and constant symbol of the theory in terms of the independently interpreted vocabulary, then every model-theoretic structure for the independently interpreted vocabulary determines a unique structure for the theory. Typically, however, correspondence rules do not provide explicit definitions for the vocabulary of a theory, in which case the independently interpreted vocabulary only determines a class of structures for the theory. In this case, the theory is called *partially interpreted*.

If the received view of theories is combined with *semantic empiricism*, the independently interpreted vocabulary is assumed to be observational and thus can be interpreted by ostension and some complementation for large sets. Accordingly, the independently interpreted vocabulary may be vague, leading to a vague interpretation of the theory even when every symbol in the theory's vocabulary can be explicitly defined in the independently interpreted vocabulary. If the independently interpreted vocabulary is only informally interpreted, the theory itself is informally interpreted by applying the informal rules of interpretation to the correspondence rules.

Semantic Theories

A theory can also be described semantically with model-theoretic structures (or, essentially equivalently, set-theoretic structures), which are then called the *models of the theory*. For semantically described theories, one typically speaks of *representations* rather than *interpretations*. There is an extensive literature on how to define representation—that is, how to distinguish relations that are representations from relations that are not. The following paragraph focuses on how to describe representations and thus assumes that a representation is already identified as such.

The representation relation between the models of the theory and the world may be an informal one, for instance, a similarity relation. More often the interpretation relation is assumed to be formal, a mapping from the models of the theory to other model-theoretic structures. As for syntactic theories, these other model-theoretic structures either in turn have to be connected to the world through some interpretation or have to be worldly structures. In the simplest case, the domain contains only objects of the world, each set corresponds to a property of or relation between objects in the world, and each function holds between objects of the world. The representation relation can be simply one of isomorphism: For each (or some) model of the theory, there is a bijection (a one-to-one and onto mapping) from its domain to the objects in the world so that each element of each set of the model of the theory is mapped to a unique element in a corresponding set of the structure of the world. Weaker mappings are possible, for instance, mappings that are not one-to-one or not on to.

Like syntactic theories, semantic theories cannot fix their own intended representations. The reason is analogous to the model-theoretic argument and can be illustrated particularly simply under the assumption that the representation relation is an isomorphism. Then a theory correctly represents the relevant aspect of the world if and only if the model of the theory is isomorphic to the structure W of the relevant aspects of the world. But again, for any set S with the same number of elements as the domain of W , there is a bijection from the domain of W to S , and a corresponding structure V with S as its domain that

is isomorphic to W . Then the model of the theory is also isomorphic to V . As for syntactic theories, then, every element of the semantic theory can represent any object whatsoever. For this reason, the intended representation of a theory needs to be fixed at least to some extent without recourse to model theory, for instance, by simply pointing to the objects that a set in the structure of the theory is intended to represent.

Informal Theories

If a theory is not completely formalized, be it in predicate logic or model theory, it can be first formalized in one of these ways and then formally interpreted as discussed earlier in this entry. But the interpretation of an informal theory can also proceed directly via the interpretation of the natural language in which it is formulated and therefore fall into the purview of linguistics and philosophy of language. There, the question of the definition of interpretation is also extensively discussed. One central assumption of this discussion is that interpretation is an intentional concept, and specifically dependent on the intentions of the person using the interpretation to, for instance, communicate the content of a theory to someone else. This assumption of intentionality occurs analogously in the debate on the definition of representation.

Restricting Interpretations

Interpretations of empirical theories can go beyond the theory proper in restricting the way the world could be so that one theory can have different interpretations. In this case, if the theory were true, according to one interpretation the world would be different from the way the world would be according to another interpretation. Such competing interpretations exist, for instance, for quantum mechanics.

Sebastian Lutz

See also Realism in Mathematics; Received View of Theories; Semantics (Introduction to Theory); Semantics (Scientific and Empirical); Syntax (Introduction to Theory); Theoretical Vocabulary; Theories, Semantic Conception of; Theories, Syntactic Conception of

Further Readings

- Andreas, H. (2017). Theoretical terms in science. In E. N. Zalta (Ed.), *The Stanford encyclopedia of philosophy* (Fall ed.). The Metaphysics Research Lab, Center for the Study of Language and Information, Stanford University. <https://plato.stanford.edu/archives/fa112017/entries/theoretical-terms-science/>
- Carnap, R. (1939). *Foundations of logic and mathematics. Foundations of the unity of science: Toward an international encyclopedia of unified science*. University of Chicago Press.
- Frigg, R., & Nguyen, J. (2020). Scientific representation. In E. N. Zalta (Ed.), *The Stanford encyclopedia of philosophy* (Spring ed.). The Metaphysics Research Lab, Center for the Study of Language and Information, Stanford University. <https://plato.stanford.edu/archives/spr2020/entries/scientific-representation/>
- Hodges, W. (1993). *Model theory*. Cambridge University Press.
- Linnebo, Ø. (2018). Platonism in the philosophy of mathematics. In E. N. Zalta (Ed.), *The Stanford encyclopedia of philosophy* (Spring ed.). The Metaphysics Research Lab, Center for the Study of Language and Information, Stanford University. <https://plato.stanford.edu/archives/spr2018/entries/platonism-mathematics/>
- Lutz, S. (2014). What's right with a syntactic approach to theories and models? *Erkenntnis*, 79(Suppl. 8), 1475–1492. <https://doi.org/10.1007/s10670-013-9578-5>
- Popper, K. R. (1968). Epistemology without a knowing subject. In B. van Rootselaar & J. Staal (Eds.), *Logic, methodology and philosophy of science III* (pp. 333–373). North-Holland.
- Przełęcki, M. (1969). *The logic of empirical theories*. Routledge & Kegan Paul/Humanities Press.
- Putnam, H. (1977). Realism and reason. *Proceedings and Addresses of the American Philosophical Association*, 50(6), 483–498. <https://doi.org/10.2307/3129784>
- Ruetsche, L. (2011). *Interpreting quantum theories: The art of the possible*. Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780199535408.001.0001>
- Tarski, A. (1944). The semantic conception of truth and the foundations of semantics. *Philosophy and Phenomenological Research*, 4(3), 341–376. <https://doi.org/10.2307/2102968>
- van Fraassen, B. C. (2008). *Scientific representation: Paradoxes of perspective*. Clarendon Press. <https://doi.org/10.1093/acprof:oso/9780199278220.001.0001>

INTUITION

Intuiting can be conceptualized as a way of direct knowing, that is, knowing “without any use of conscious reasoning” (Sinclair & Ashkanasy, 2005, p. 357), through a process that seems to bypass sequential (i.e., step-by-step) reasoning. Although the terms *intuiting* and *intuition* are often used interchangeably, it is important to differentiate between them. Whereas intuiting is the process of *direct knowing*, intuition is the outcome of this process. Top professionals in any area of expertise, including scientists, engineers, business managers, musicians, chefs, athletes, and designers, are esteemed precisely for their *sense of the game* in their domains, even when formally they are recognized for performance. Yet, until recently, it was mainly philosophers who advocated for intuition, while the scientific community was wary of it.

What Is Intuition?

Nowadays, numerous scientists admit to relying on intuition as a part of scientific inquiry. In fact, based on interviews with 20 top scientists, including 17 Nobel laureates, Viktor Dörfler and Colin Eden (2019: 538) assert that “no significant research result seems to have been achieved without intuition playing a major role in the process.” The interviewed Nobel laureates offered examples of using intuition in some form in their work, such as their ability to see simple patterns in complex situations. These examples, while not necessarily labeled as intuition, fit the description of direct knowing without sequential reasoning or understanding how the knowledge was obtained.

Scholars generally agree about recognizing patterns with which an intuitor had previous experience; this is often referred to as *expert intuition*. Some scholars, however, argue that experts may also identify previously unseen patterns and, further, can create new patterns that did not exist before, which is the basis of creative intuition.

Contemporary scholars agree that people tend to go back and forth between intuitive and nonintuitive modes of working through things, a thesis that was proposed decades ago by both Michael Polanyi and Henri Bergson, two

scientists-turned-philosophers who advocated for intuition. Intuition frequently endures unfounded criticism of being recklessly impulsive and ignoring facts. Contrary to this misrepresentation of intuition, research shows that scientists, much like investors, typically diligently analyze available data, as well as use intuition to sense the validity of their assessments. Intuiting is far from acting on a whim; it is an elaborate, highly complex process that incorporates facts and processes them in a holistic way. Perhaps the very reason for the non-conscious nature of intuiting is that the conscious mind is not always capable of the level of complexity often required for situations at hand. Even in highly quantitative domains, such as experimental science, software engineering, and finance, intuitive decision makers frequently use stories rather than computations to frame their decisions. Gary Klein describes intuitive decisions taken by firefighters: They take calculated risks rooted in elaborate perceptions. Laura Huang and Jone L. Pearce, based on their study of angel investors, suggest that decision makers who use intuition in their practice, are less likely to miss seemingly trivial details that are in conflict with analytical presumptions about a situation; these details often provide essential clues about investment outcomes. A number of prominent intuition researchers approached intuition by exploring some of its components that can allegedly be measured, while others moved on to exploring the sensory, experiential aspects of intuition. Intuition is less of a construct to be measured and quantified, and more of a phenomenon to be observed, sensed, and synced with in order to understand its nature.

What Intuition Is Not

When dealing with something new, one way to explain it is through what it is not, demonstrating how this new entity is not like the known entities; initially, there is a lack of vocabulary to explain the new entity. In line with this logic, it is important to point out that intuition is *not random guessing*. In guessing, none of the solutions is recognizable, but during intuiting, some solutions may produce psychophysiological coherence or an adverse response. If the response is accurately interpreted, it can be useful in selecting a preferred option. Importantly, intuition is correct far more

often than the probability of random guessing. Intuition is not instinct; unlike *instinct*, which is hardwired and is always unconscious, the quality of intuition can improve as one's experience increases. Intuition is also not *heuristics*. In contrast to heuristics, which means applying a pre-defined rule to a specific situation, intuition is adaptive and may involve recognition of patterns not seen before. Related to this, intuition is also unlike *biases*, which are tacitly held preferences that manifest in specific decision situations. This does not mean that intuition is never biased, since complex pattern recognition may involve biases, as shown by Amos Tversky and Daniel Kahneman (1974), but sequential reasoning can be equally subject to biases. Intuition is more than just a *gut feel*, although it is often described as something felt at the pit of one's stomach; intuiting can involve one's whole body. Jennifer Frank Tania proposes that intuition is not simply a gut feel, but rather, a whole-body experience, extending into space around the subject's body as well as into the imaginary realm.

Further, intuition is not a *sixth sense*, as it is often described. Intuition seems to make use of all of our senses and capacities. Intuiting is not *super-natural*; what is understood about intuition can be explained in terms of people's natural abilities to perceive themselves and their environments. Neither is intuition *extrasensory*, as it relies on the body's psychophysiological system, including senses of vision, hearing, taste, smell, and touch, in addition to other senses, to receive information or make sense of something without resorting to sequential analysis. Some discomfort with the subject of intuition may be due to its association with underresearched and underexplained phenomena like remote viewing, telepathy (e.g., mind-to-mind communication), and precognition (sensing future events). Although studies related to these topics still remain marginalized, broader intuition studies and projects are gaining acceptance, ranging from getting published in top journals like *Human Relations* to being explored by the U.S. military.

Intuition is often contrasted with analysis, which is incorrect. The opposite of analysis, which dissects something to examine its parts, is *synthesis*, which fuses separate parts into a larger phenomenon to understand a bigger picture. Although intuition can involve synthesis, the two are not the same.

Intuitive Judgment and Intuitive Insight

In the management literature on intuition, there appears to be an underlying assumption that all intuitions are judgments. This neglects a possibility of intuitive insight, the "*aha* moment" resulting from intuiting rather than from deliberate sequential reasoning. Dörfler and Fran Ackermann put forth a notion of *intuitive insight*, suggesting that insights can be both intuitive and nonintuitive, much like judgments that can be intuitive and nonintuitive. Often both intuitive insight and intuitive judgment appear in addressing a problem. For example, a decision maker can create new solutions through intuitive insights and then judge alternative solutions intuitively or otherwise. In another example, expert chefs rely on both intuitive judgment and intuitive insight to consistently produce creative, original, interesting food experiences; chefs may use intuitive insight to envision new dishes, and intuitive judgment to decide whether these dishes may work out in the context of their restaurants.

Features of Intuiting and Intuition

Currently, it seems that intuition research is being informed by practice more than the other way around; the most useful and relevant data on intuition are gathered through interviewing, observing, and surveying expert practitioners. For instance, it was primarily through interviewing decision makers and creative problem solvers that intuition researchers reached a consensus (for the time being) on six necessary features of the process of intuiting and its outcome, intuition:

- Intuiting is *rapid*, meaning that it occurs almost instantaneously, and in this respect, it is similar to guessing. However, it is guessing which is *frequently correct* (in line with Simon, 1983, p. 25).
- Intuiting is *logical*, meaning that it operates independently of the general principles of reasoning that we call logic; therefore it is neither *logical* nor *illogical*, as it neither follows nor contradicts the rules of logic. In deliberate sequential reasoning, we work our way through the logical connections from a particular

problem to the idea that addresses it, while intuiting leaps to the solution.

- Intuiting is *tacit* in the sense that intuitors can describe the outcome but not the way they arrived at it. Although intuitors are often good at defending their intuitive judgments or insights, there is no evidence that the justification has anything to do with the way the intuiting was used. The tacit quality of intuiting suggests that the process of intuiting takes places outside the scope of the conscious mind, which is why it is also referred to as *nonconscious*.
- Intuition is *holistic* or *gestalt*, as it is about the *big picture*, including broad context, and far-reaching implications. It even takes into account items inaccessible to deliberate sequential reasoning, sometimes referred to as the *unknown unknowns*. Furthermore, intuiting also involves the totality of the intuitor's experiences.
- Intuition, as the outcome of intuiting, has an *intrinsic certainty*: Intuitors are confident they have the right answer (or one right answer). Even though intuition often brings with it a *feeling of knowing*, this does not make intuiting infallible, and intuitors do not usually claim infallibility. Robert A. Burton (2009, p. 101) points out that the *feeling of knowing* can occur with both proven and unproven insights and is not a reliable indicator of accuracy.
- Finally, intuition, as the outcome of intuiting, is considered to be *spontaneous*. This means that intuiting does not require conscious effort at the moment when it happens. This is the one feature on which intuition scholars (academics who study intuition empirically) and intuition practitioners (those who tap into their intuition in their practice) disagree. The spontaneous nature of intuiting has recently been brought into question, as intuition practitioners have been reporting intentional summoning and use of intuition. The two contradictory views may be reconciled by allowing for the *spontaneous nature* of intuition, that is, useful and relevant information emerging into consciousness without deliberate effort, and the *intentional use of intuiting*, in other words a process of deliberately conjuring the process of intuiting on cue.

Besides these six features of intuiting, it is also often mentioned that intuiting is frequently

accompanied by somatic (visceral) effects and affective charges. Whereas the role of affect in intuiting and intuition has been addressed, the somatic aspects are much less understood. It seems inevitable that some aspect of somatic experience will soon be added to the list of the aforementioned six characteristics.

Facets of Intuitive Awareness

The phrase *facets of intuiting* refers to the various pathways through which intuitive messages arrive or become known to an individual. For the physical facet of intuiting, Frances Vaughan suggests that one may notice muscle tension, stomachache, or back pain seemingly not in response to anything in the environment of which the person is aware (1989, p. 47). This lack of congruence between somatic experience and either the environment or one's mental state may prompt one to examine more closely what is going on beyond the obvious.

Vaughan clusters together both emotions and feelings in the emotional facet of intuiting. For instance, one may feel safe in some place or in presence of a particular person, or to the contrary, one may experience a sudden emergence of anger with a person, seemingly without reason. Vaughan suggests learning to recognize our emotional states, and notice how one feels, as these feelings may affect perceptions, memories, and interpretations. There is some disagreement among scholars regarding the role of affect in intuiting: whether it always accompanies intuition and whether its presence is prognostic or reactive. Marta Sinclair noticed that fluctuation of emotions can serve as a channel for intuiting rather than being an add-on to intuition. Huang and Pearce suggest that affect is a reaction to the level of decision maker's confidence in the decision.

The mental facet of intuiting is related to thinking: It could be an intuitive insight for a challenge that one is dealing with, or an unexpected thought, seemingly without a prompt from the environment, and not resulting from sequential thinking through ideas. This facet of intuiting is the most studied and the best understood, as it is the mental facet that provides account of both intuitive judgments and insights.

The spiritual facet of intuiting, according to Vaughan, is a transcendent, holistic form of

intuition. While there is no clear description of spiritual intuition, it probably closest to the notion of indwelling, that is, attending to the object of one's interest, much like being fully engaged in listening to a musical composition or exploring a work of art.

Interplay Between Intuition and Sequential Reasoning

Considerable gaps remain in our understanding of the dynamic between deliberate sequential reasoning and intuition. One concern that academics, practitioners, and corporate leaders have about employing intuition is that if intuiting becomes widely accepted as a valid way of knowing, students and employees may be tempted to use it as a substitute for rigorous analysis, logical argument, or deep thinking. While intuiting cannot replace all aspects of deliberate sequential reasoning, thinking and data crunching cannot replace intuition for types of problems that require a sensory or experiential component, like creating psychological safety in a team.

Until fairly recently, it was unclear whether intuition and deliberate sequential reasoning are two aspects of one continuum, or whether these are two separate dimensions, as the dual-process theories suggest. Early ideas of the interplay between intuition and analysis suggested a continuum, on which one would be either more or less intuitive. As more recent research indicates, intuition and deliberate sequential reasoning are likely *two separate dimensions*. One can be good or bad at both, or better at one than the other, relying on each dimension more or less, depending on one's level of expertise, the type of problem at hand, and the environment. Expert decision makers, scientists, design engineers, and other practitioners reconcile their intuition and formal reasoning by creating a cohesive narrative.

Factors Affecting Intuition

The quality of intuition is related to three factors: characteristics of the intuitor, the problem at hand, and the environment. Level of expertise is one of the most frequently discussed characteristics of an intuitor. "Intuiting becomes the dominant mode of knowing at the highest level of expertise" (Dörfler & Stierand, 2017, p. 4). A closer look at studies claiming to have provided

experimental evidence on the failure of intuition reveals that these experiments targeted intuition of novices. In contrast, studies that have found empirical evidence of intuition working well, typically focused on intuition of people at a high level of expertise (predominantly using nonexperimental methods, such as interviews and observations in natural context). Some researchers point out that a high level of expertise is necessary for good intuitions. Therefore, it seems that expertise fuels intuiting, enabling intuiting to become the dominant form of knowing.

In addition to expertise, other factors that influence intuiting on an individual level are attitude toward intuition (dismissive vs. accepting); physical, emotional, and intellectual self-awareness; and preference for intuitive versus sequential processing style.

Characteristics of problems and circumstances under which intuiting is the dominant mode of reasoning are typically those that do not allow the use of deliberate sequential reasoning, including uncertainty, limited data, no precedent, time pressure, volatility of the situation, high complexity, and ambiguity—in other words, the usual environment of today's organizations.

An organization's openness to intuiting and a cultural fit for use of intuition contribute to employees' willingness to employ (and disclose the use of) intuiting. Accountability to others seems to reduce the inclination of knowledge workers to rely on intuition. Organizational culture against intuiting as a way of knowing may lead to suppressing intuitions in an effort to conform or to get approval for the way the answers have been arrived at. However, the danger of downplaying or hiding the use of intuiting is in reducing transparency around the problem-solving process.

Intuition as Sensing Plus Sensemaking

Practice-based explanation of intuiting is founded on accounts of *intuition practitioners*—practitioners who have expertise in the use of intuiting and who consistently use intuiting in their practice, intentionally and on cue. Intuition practitioners see intuiting as a sensory-based *direct way of obtaining awareness* of phenomena. *Affect and visceral sensations* are essential to intuition practitioners: affective and visceral cues typically

serve as starting points that make intuition practitioners pay attention to intuiting. Intuition practitioners employ a sensory-based way of obtaining awareness of phenomena, achieving firsthand experience. For the purpose of this entry, we use the term *sensory* to mean “composed of the five primary senses, as well as visceral (e.g., hunger), affective (e.g., love), and mental sensations (e.g., pride).”

Although intuition practitioners may or may not describe intuiting as rapid, tacit, holistic, alogical, and having intrinsic certainty, as scholars do, they often talk about intuition as *both spontaneous and intentional*. For example, intuiting could be engaged by asking oneself a question about a situation, or defining a desired outcome, and staying attuned with soft focus to sensory and environmental cues emerging in response to the initial inquiry. After posing a question, practitioners may attend to shifts in their mental, physical, and affective states. These shifts seem to occur without apparent causes either from the external environment or from internal processes and hence are labeled as intuitive.

Someone not experienced with paying attention to sensing may wonder how it is possible to get intuitive information seemingly out of nowhere unless there is a way to justify it as expertise. Intuition practitioners suggest that, over time, a student of intuiting could learn to understand various sensations underpinning their intuiting. Consciously making sense of these sensations can create a useful, actionable narrative. As an analogy, to a musically challenged person, observing someone differentiate between F# and G in the third octave just by hearing the sounds may seem like a lucky guess or a sort of a magic trick. Similarly, when it comes to intuiting, it may seem just as mysterious to observe someone differentiate among the meanings of their sensory signals. To highly trained musicians, even the notion of confusing F# and G may seem strange because they perceive the two notes as two distinct sounds. Similarly, someone who is practiced in intuition is able to distinguish between an intuition about a positive resolution versus unfounded hope.

The academic literature on intuition treats intuition as a unitary construct. A more nuanced way to look at intuiting is in terms of a process composed of two phases: *sensing* and *sensemaking*,

resulting in intuitive judgment or intuitive insight. Expert practitioners of intuition who admit to using intuition in their work actively engage with both sensing and sensemaking.

The sensing phase of intuiting delivers signals which an attentive person can perceive as visual images, auditory sensations, phrases, and changes in physical and emotional state. Intuition arrives to each person in a language or form that is unique to that person, therefore it cannot be standardized. For example, the same visual image of a dog can represent protection to one person and fear to another, or it can mean different things to the same person in different contexts. Hence, if sensing is not followed by sensemaking, the product of sensing cannot be utilized for intuiting on cue. Sensemaking is concerned with revealing the meaning of sensed signals. The sense making phase appears to be tied to one's domain of expertise relevant to the particular situation (i.e., expert intuition). The pitfall of expert intuition is that experts may not be aware of the sensing part of intuiting, as they only perceive the signals of which they have already made sense. Thus, they do not distinguish between the *outcome of intuiting* and the *outcome of sensemaking*, being oblivious to the notion that intuiting as a whole includes not only sensemaking but also sensing. This may lead to a missed opportunity to improve intuiting by developing one's sensory awareness.

Conclusion

Knowledge workers in today's organizational environment must be prepared to operate in volatile, uncertain, complex, and ambiguous conditions. Under these conditions, best or satisfactory solutions for complex problems arise through cognitively flexible response rather than additional analysis; scientists, managers, and design engineers, for example, may need to use their intuition (knowing without knowing how) as they move through uncertainty. How can analytically minded practitioners become better at intuiting? They need to develop attention to *sensing* through enhanced awareness of their physical, emotional, and mental states, and they should advance their subject area expertise to improve *sensemaking*. At the highest level of expertise, it seems that

intuiting, supported by deliberation, is the predominant way of working through a problem. Yet, while Western education places great emphasis on teaching sequential reasoning skills as means to achieve expertise, it often neglects the development of sensory capacities, which appear to be essential for exceptional performance.

Viktor Dörfler and Alina Bes

See also Analysis; Artificial Intelligence; Cognitive Science; Knowledge; Philosophy of Mind; Rationality; Reasoning

Further Readings

- Bas, A., Dörfler, V., & Sinclair, M. (2019). *Intuiting process as sensing plus sensemaking* [Paper presentation]. AoM 2019: 79th Annual Meeting of the Academy of Management. <http://aom.org/annualmeeting/>
- Bergson, H. (1911). *Creative evolution*. Henry Holt and Company.
- Burton, R. A. (2009). *On being certain: Believing you are right even when you're not*. St. Martin's Press.
- Dörfler, V., & Ackermann, F. (2012). Understanding intuition: The case for two forms of intuition. *Management Learning*, 43(5), 545–564. <https://doi.org/10.1177/1350507611434686>
- Dörfler, V., & Bas, A. (2020). Intuition: Scientific, non-scientific or unscientific? In M. Sinclair (Ed.), *Handbook of intuition research as practice* (pp. 293–305). Edward Elgar.
- Dörfler, V., & Eden, C. (2019). Understanding “Expert” scientists: Implications for management and organization research. *Management Learning*, 50(5), 534–555. <https://doi.org/10.1177/1350507619866652>
- Dörfler, V., & Stierand, M. (2017). The underpinnings of intuition. In J. Liebowitz, J. Paliszkievicz, & J. Goluchowski (Eds.), *Intuition, trust, and analytics* (pp. 3–20). Taylor & Francis.
- Huang, L., & Pearce, J. L. (2015). Managing the unknowable: The effectiveness of early-stage investor gut feel in entrepreneurial investment decisions. *Administrative Science Quarterly*, 60(4), 634–670.
- Polányi, M. (1962/2002). *Personal knowledge: Towards a post-critical philosophy*. Routledge.
- Simon, H. A. (1983). *Reason in human affairs*. Stanford University Press.
- Sinclair, M., & Ashkanasy, N. M. (2005). Intuition: Myth or a decision-making tool? *Management Learning*, 36(3), 353–370. <https://doi.org/10.1177/1350507605055351>
- Tantia, J. F. (2014). Is intuition embodied? A phenomenological study of clinical intuition in somatic psychotherapy practice. *Body, Movement and Dance in Psychotherapy*, 9(4), 211–223. <https://doi.org/10.1080/17432979.2014.931888>
- Tversky, A., & Kahneman, D. (1974, September 27). Judgment under uncertainty: Heuristics and biases. *Science*, 185(4157), 1124–1131. <https://doi.org/10.1126/science.185.4157.1124>
- Vaughan, F. E. (1989). Varieties of intuitive experience. In W. H. Agor (Ed.), *Intuition in organizations* (pp. 40–61). SAGE Publications.

INTUITIONISM IN LOGIC AND MATHEMATICS

The term *intuitionism* derives from the 18th-century philosopher Immanuel Kant's account of temporal intuition as being a priori. Intuitionism considers mathematics purely the result of constructive mental activity and as having no principles relating to an objective reality. On this view, logic and mathematics are not activities that reveal any deep levels of reality but are simply the application of internal methods to build more complex mental constructs.

Intuitionism is characterized by its interpretation of what it means for a mathematical or logical statement to be true. In mathematical intuitionism, a mathematical statement corresponds to a mental construction, and a mathematician can only assert the truth of a statement by verifying its validity through intuition; that is, the truth of a statement is equivalent to the mathematician's being able to *intuit* the statement, something that does not of necessity relate to its provability. Communication between mathematicians serves only as a means to create the same mental processes in different minds.

A mathematical object is considered the product of a mind so that if it can be constructed then it exists. This entry first explains the foundations of intuitionism, its relation to the concepts of infinity and natural numbers, and continues with a history of intuitionism including the contributions of its key thinkers up to the present day. Following a section outlining the practice of intuitionistic logic, the entry concludes with a description of recent and current developments in the field.

Forms of Intuitionism

Intuitionism is a form of mathematical constructivism since it asserts that it is necessary to find (construct) a mathematical object to prove it exists. Put another way, to claim that an object with certain properties exists is to claim that an object with certain properties can be constructed. A mathematical object is considered a construct of the mind; therefore, its existence is equivalent to the possibility of its construction. This is in contrast with the classical approach, which tells us that existence of an entity can be proved by refuting its nonexistence. Intuitionism disputes this, stating that refuting an object's nonexistence does not prove that it is possible to find a construction for the apparently valid object, something that is required in order to assert its existence.

Time is an essential factor in dependent on the notion that statements can become provable in time and could therefore become intuitionistically valid where they were not so before.

Intuitionism and Infinity

The various forms of intuitionism feature differing positions on the meaning and reality of infinity. The term *potential infinity* refers to a mathematical procedure that has an unending series of steps. Once each step is completed there is always another step to be performed. The process of counting 1, 2, 3, and so on is an example. *Actual infinity* refers to a complete mathematical object that contains an infinite number of elements. The set of natural numbers $\{0, 1, 2, 3, 4, 5, \dots\}$ is an example.

According to the German mathematician Georg Cantor's set theory, there are many different infinite sets, and some are larger than others. For example, the set of all the real numbers, $\mathbb{R}$, is larger than $\mathbb{N}$, the set of all natural numbers, because all attempts to put natural numbers into one-to-one correspondence with real numbers will fail, and there will always be an infinite number of real numbers left over. Any infinite set that can be placed in one-to-one correspondence with the natural numbers is said to be *countable*, while infinite sets larger than this are said to be *uncountable*. This approach formed the foundations of mainstream mathematics.

Intuitionism set out to reject the reality of uncountable infinite sets. The Dutch mathematician L. E. J. Brouwer rejected the concept of *actual infinity* but accepted the idea of *potential infinity*. He claimed that there is no evidence to support the belief in the existential character of the totality of all natural numbers. The sequence of numbers that grows beyond any stage already reached by passing to the next number is a manifold of possibilities open toward infinity. While it remains forever in the status of creation it is not a closed realm of things existing in themselves. Brouwer insisted that mathematics in its classical form went beyond all human possibilities of realization and beyond statements that can claim meaning and truth founded on evidence.

Finitism

Finitism is an extreme form of intuitionism that rejects the idea of potential infinity. Finitism holds that mathematical objects cannot exist unless they can be constructed from the natural numbers in a finite number of steps. The Platonic idea of mathematical objects such as infinite sets being accepted as legitimate mathematical objects is rejected. Natural numbers are accepted as existing, but the set of all natural numbers is not considered to exist; therefore, discussion of infinite domains is not meaningful. In *The Philosophy of Set Theory*, Mary Tiles terms those who allow for *countably* infinite objects as *classical finitists* and those who disallow *countably* infinite objects as *strict finitists*. Leopold Kronecker was a famous finitist who believed that God created all the natural numbers, but everything else was a human creation.

The use of infinite objects has long been a controversial subject in mathematics. Brouwer rejected the existence of infinite objects until they are constructed. David Hilbert believed that finite mathematical objects are concrete objects and infinite mathematical objects are ideal objects. He believed it is possible to illustrate that any theorem about finite mathematical objects obtained using ideal infinite objects can also be obtained without them. Allowing that there are infinite mathematical objects does not cause a problem with finite objects.

Origins and History of Intuitionism

Intuitionism can be traced back to the mathematics of antiquity, then later to academics including Carl Gauss, Leopold Kronecker, Henri Poincaré, Henri Lebesgue, and Émile Borel. During the 19th century, discussions between German mathematicians Georg Cantor and his tutor Kronecker were instrumental to the development of intuitionism, later followed by Gottlob Frege and the critiques of his work by Bertrand Russell.

At the beginning of the 20th century, Brouwer advanced his comprehensive critique of classical mathematics, a program we now know as intuitionism. Brouwer systematically constructed the principal branches of intuitionist mathematics including set theory, mathematical analysis, geometry, and topology.

French philosopher Henri Bergson developed a version of intuitionism in his *Introduction to Metaphysics*, where he posited that an object can be known in two distinct ways, either *absolutely* or *relatively*. Knowledge is attainable relatively through analysis and absolutely through intuition. Bergson explains intuition using the analogy of a city. We can only grasp the dimensional value of a city by walking through it, we can never know it in the same way, even if we have photographs taken from every possible point of view. Similarly, the experience of reading a line of Homer: While we can translate the line and heap commentary upon it, our commentary can never grasp the simple dimensional value of experiencing the poem in its originality. Intuition, for Bergson, is the act of getting back to the things themselves.

Georg Cantor was the inventor of transfinite arithmetic, which was subsequently rejected by many esteemed mathematicians, including his tutor Kronecker—a confirmed finitist. Cantor's logical methods to prove his results in transfinite arithmetic are essentially the same as those later used by Russell in constructing his paradox. Cantor was responsible for the creation of the first version of set theory at the end of the 19th century as part of his study on infinite sets although it was Frege's attempts to reduce all mathematics to a logical formulation via set theory that led to Russell's paradox. Frege's axiomatic set theory was developed in response to early attempts at set

theory, with the goal of determining precisely what operations were allowed and when.

This was then derailed by Bertrand Russell who discovered Russell's paradox, demonstrating that one of Frege's rules of self-reference was self-contradictory. Russell discovered that naïve set theory contained a contradiction. It had been thought that for every concept there exists a set of things that fall under that concept. For example, corresponding to the set *feather* are all the feathers in the world. Even concepts such as *mermaid* are associated with sets, namely, the empty set. Russell pointed out, however, that there is no set corresponding to the set *not a member of itself*.

If there had been such a set—a set of all the sets that are not members of themselves—and we called this set *S*, then is *S* a member of itself?

If it is, then it is not (because all the sets in *S* are not members of themselves), and if *S* is not a member of itself, then it is (because all the sets not in *S* are members of themselves). Either way a contradiction ensues and therefore we can safely say that there is no such set as *S*.

Kronecker took the natural numbers as primitively generable and primitively graspable mental constructions that can subsequently construct other kinds of mathematical objects. Neither the natural numbers nor any other infinitely numerous objects of any kind can be thought of as comprising a completed totality as an object to be treated in its own right. Mathematical objects, since they are mental constructions, depend on their mode of construction, presentation, or description for their identity. They possess intentional identity conditions, which means that their identity relation is effectively decidable. By contrast, the extensional identity in mathematical objects is not generally decidable.

Stephen Kleene subsequently submitted a more rational account of intuitionism in his *Introduction to Meta-Mathematics* in 1952. Kleene helped to lay the foundations for theoretical computer science and was a founder of the branch of mathematical logic known as *recursion theory*.

Early in the 20th century, Brouwer gave the first formal exposition of intuitionism, when he set out his criticisms of certain conceptions in classical mathematics. He questioned the status of existence in mathematics, in terms of how we may

consider the established existence of an infinite set. Brouwer took the notion of natural number to arise out of our a priori conception of successive moments in time. In contrast with Platonism, which holds that mathematical concepts exist independent of any human realization of them, Brouwer held that only those mathematical concepts that can be demonstrated, or constructed, following a finite number of steps are legitimate. Unlike many other mathematicians, Brouwer was willing to abandon the vast realms of mathematics built on nonconstructive proofs.

Brouwer believed mathematics to be a language-less creation of the mind. For him, time is the only a priori notion as proposed by Kant. Brouwer distinguished two acts of intuitionism: The first act was to completely separate mathematics from mathematical language and from theoretical logic. Intuitionistic mathematics is an activity originating in the perception of a movement of time. This act gives rise to the natural numbers and restricts the kind of reasoning we may do, including that of the principle of the *excluded middle*. Brouwer's view differed from other constructive views in allowing for the construction of infinite sequences and that mathematics is the creation of a free mind, making it neither Platonic nor formal.

The second act is the creation of a *free choice sequence*. This is given by a constructing agent or creative subject who is free, at any stage of the progression of a sequence, to choose the next number in the sequence. For example, if someone is asked to produce a binary free choice sequence they might start "000. . .0" for a hundred digits then completely unexpectedly choose "1" as the next digit. A real number classically thought of as a converging sequence does not need to be given a determinate rule. It is only subject to a Cauchy restriction or a certain rate of convergence.

So, for Brouwer in intuitionism in mathematics the free creation of the human mind is at work, and objects only exist insofar as they can be mentally constituted. The natural numbers and the continuum are basic notions given to us directly by intuition. His intuitionistic theory of mathematics differs from that of others on the basis of the free choice sequence.

Criticism of the Excluded Middle

Brouwer criticized the law of the excluded middle. He became convinced that universal validity of contradiction proofs for existence proofs were unwarranted. This has significant implications for mathematics because it is performed on an everyday basis. To say "A or B" to an intuitionist is to claim that either "A or B" can be proved. The law of the excluded middle, that is to say "A or not A," cannot be allowed as a general inference rule.

The principle of the excluded middle asserts that for every proposition either that proposition is true or its negation is true. The excluded middle has its earliest formulation in Aristotle's principle of noncontradiction, wherein he states that of two contradictory propositions (where one proposition is the negation of the other) one must be true and the other false. The principle posits that it is impossible for there to be anything *between* the two parts of a contradiction. For example, where "P" proposes that

Socrates is mortal

then the "law of excluded middle" affirms that

Either Socrates is mortal, or it is not the case that Socrates is mortal.

The potential *middle* position (that Socrates is neither mortal nor nonmortal) is logically disallowed so that the first possibility whereby Socrates is mortal, or the negation of this possibility (it is not the case that Socrates is mortal) must be true.

An intuitionist would not accept this argument without further support for that statement since the proof is nonconstructivist. Negation as a concept has a different interpretation in intuitionist logic compared with classical logic. In classical logic negating, a statement involves asserting that a statement is false. An intuitionist on the other hand would take this as meaning that the statement is refutable, in other words, that there is a counterexample (a proof there is *no proof*). There appears then to be an asymmetry between positive and negative statements in intuitionism: for example, if statement P is provable, it is clearly impossible to prove that there is "no proof of P." However, even if we can show that no *disproof* of P is possible, we cannot conclude from its absence that there is a proof of P. P is thus a stronger

statement than not-not-P. To reiterate, just because there is no proof that there is “no proof of P,” we cannot conclude that there “is a proof of P.”

One way to think of this odd consequence of the law of the excluded middle is that when it is projected beyond the immediate circumstances of our ability to ascertain its truth by observation it becomes problematic. On this basis, intuitionists must reject it. Extrapolating the law of the excluded middle into the future or into the past brings with it difficulties. Our strong intuitive sense of realism means it doesn't make it feel strange to assert that something that hasn't happened yet either will happen or will not happen. There is no third possibility, so one of the propositions is true and the other false; just because they haven't happened yet does not make them false. This remains a problem for realists. However, the risk of taking this view is that we can end up in determinism. Assuming that all future contingents have an absolute truth implies they are set in stone and there is no freedom of choice, an orientation that many would reject.

As an example, take Goldbach's conjecture that every even number is the sum of two primes. As yet, no one has discovered either a proof or a counterexample. From a realist perspective, it makes sense to suppose that this conjecture might be true since every one of the infinite series of even numbers is a sum of two primes (although there may be no proof to be discovered).

For an intuitionist, the only thing that could make it true that all even numbers are the sum of two primes is that there is a *proof*. From the intuitionist point of view, there may be no proof and no counterexample and thus nothing to give the conjecture a truth-value.

Intuitionistic Logic

Mathematics requires intuitionistic logic, or *constructivist* logic, to preserve its justification, rather than truth. Intuitionistic logic, also known as *constructivist logic*, is the symbolic logical system developed by the Dutch mathematician Arend Heyting (1898–1980) aimed at providing a basis for Brouwer's program of intuitionism. Intuitionistic logic substitutes justification for truth in its logical calculus. The logical calculus is responsible for preserving justification rather than truth,

giving philosophical support to several schools of philosophy, notably Michael Dummett's *anti-realism*. Dummett was a highly influential philosopher who did not find Brouwer's interpretation of intuitionism convincing because it relies on the idea that a mathematical construction is a process carried out by an individual mathematician in the privacy of their own mind, hence seeming to identify the meaning that one individual attaches to a mathematical term with a private mental object, which only that person has access to.

The system preserves justification, rather than truth, across transformations yielding derived propositions. Practically, there is a good case for using intuitionistic logic, since it has the property of existence, which renders it suitable for use in other forms of mathematical constructivism. The syntax of formulae in intuitionistic logic is similar to that of propositional logic; however, its connectives are not interdefinable in the same way. In intuitionistic logic, it is possible to disprove many of the tautologies of classical logic. Examples include the law of the excluded middle, Peirce's law, and even double negation elimination. In classical logic, both and also are theorems. In intuitionistic logic, only the former is a theorem: Double negation can be introduced but cannot be eliminated. In observing that many classical valid tautologies are not theorems of intuitionistic logic, the proof theory of classical logic is weakened.

The question of how much structure sensory input has of itself and how much we give it through processing is one of the central problems of philosophy and in cognitive science in general. Intuitionism, unlike realism, holds that mathematical entities cannot exist independently of the human mind. Mathematics is not out there to *be discovered* but is rather constructed as an exercise in human intuition. Mathematical objects are created through mental activity.

Ongoing Development of Intuitionism

Although intuitionism has never replaced classical mathematics as the standard view, it has attracted much attention and continues to do so today. Philosophers and mathematicians alike were forced to acknowledge the lack of an epistemological and ontological basis of mathematics, even when their loyalties lay elsewhere. Even the likes of Kurt

Gödel, a lifelong Platonist, and Hermann Weyl, who later stopped practicing intuitionistic mathematics, continued to admire Brouwer's intuitionistic philosophy of mathematics.

In 1936, Alan Turing (1912–1954) considered nonconstructive systems of logic with which not all the steps in a proof are mechanical, some being intuitive. The Church–Turing thesis (also called *Church's thesis*) emerged from research in the foundations of logic and mathematics. The conjecture bears on what a computer ultimately can do (and thus what it cannot do). Later Stephen Cole Kleene brought forth a more rational consideration of intuitionism, where he stated that there are no paradoxes in Cantorian set theory—thereby calling into question the program of intuitionist mathematics.

In the late 1940s, a group of mathematical logicians headed by A. A. Markov undertook the task of developing a constructive mathematics, with a tendency close to that of mathematical intuitionism known as *recursive constructive mathematics* or *recursive function theory* (RUSS).

Bishop's constructive mathematics in *Foundations of Constructive Analysis* (1967) provided a constructive development of a significant part of 20th-century analysis. His mathematics (termed *BISH*), unlike intuitionistic or recursive constructive mathematics, admits of many interpretations. In fact, the results and proofs of BISH can be interpreted in most reasonable models of computable mathematics, for example, Klaus Weihrauch's type two effectivity theory and Andrej Bauer's model.

Fred Richman took logic as the primary characteristic of constructive mathematics. This was in contrast to the pioneers of constructivism Brouwer, Heyting, Markov, and Bishop who believed in the primacy of mathematics over logic. Richman saw constructive mathematics as characterized by its methodology based on the use of intuitionistic logic.

Per Martin-Löf's constructive type theory (ML) published in *Notes on Constructive Mathematics* (1968) is in the spirit of RUSS rather than BISH; it had the advantage of facilitating the extraction of programs from proofs. This led to work on the

implementation of constructive mathematics by Robert L. Constable; Susumu Hayashi and Hiroshi Nakano; and Helmut Schwichtenberg. Recent implementations of type theory for proof extraction include the programs Coq and Agda.

The field of mathematical intuitionism is still being intensively developed. The emphasis placed by intuitionists on the effectiveness of its results sits well with the computational tendency in modern mathematics and continues to draw many outstanding minds to its study.

Thora L. Fitzpatrick

See also Church–Turing Thesis; Gödel's Incompleteness Theorems; Realism in Mathematics; Set Theory

Further Readings

- Anglin, W. S. (1994). *Mathematics: A concise history and philosophy*. Springer-Verlag.
- Davis, M. (1965). *The undecidable*. Raven Press.
- Davis, M. (2000). *Engines of logic: Mathematicians and the origin of the computer*. W. W. Norton.
- Dawson, J. W. (1997). *Logical dilemmas: The life and work of Kurt Gödel*. A. K. Peters.
- Goldstein, R. (2005). *Incompleteness: The proof and paradox of Kurt Gödel*. W. W. Norton.
- Hesseling, D. E. (2003). *Gnomes in the fog. The reception of Brouwer's intuitionism in the 1920s*. Birkhäuser.
- Heyting, A. (1971). *Intuitionism: An introduction* (3rd ed.). North-Holland.
- Kleene, S. C. (1991). *Introduction to meta-mathematics*. North-Holland.
- Kleene, S. C., & Vesley, R. E. (1965). *The foundations of intuitionistic mathematics*. North-Holland.
- Reid, C. (1996). *Hilbert, Copernicus* (2nd ed.). Springer-Verlag.
- Rosenbloom, P. (1950). *The elements of mathematical logic*. Dover.
- Tiles, M. (1989). *The philosophy of set theory: An historical introduction to Cantor's paradise*. Blackwell. (Reprinted by Dover 2004)
- van Heijenoort, J. (1967). *From Frege to Gödel, a source book in mathematical logic, 1879–1931*. Harvard University Press.

J

JUSTIFICATION

Justification (J) is defined as the reason or set of reasons for which someone holds a belief. J focuses on the reasoning behind affirming a statement or proposition. J is a *rational defense* of an argument, as opposed to an *imposition* based in power, coercion, or the tradition of an argument. Entities that can be justified include from beliefs, values, emotions, or behavior to laws, judicial decision, or scientific theories. J is an account of why a statement can be considered as a true one. J is an explanation on how one knows what one knows. There is a sort of J addressed to justify statements, and there is a sort of J addressed to justify behavior. J can be based on fact (positive justifications), values (normative justifications), and inferences (logical justifications). Examples of J include trying to persuade others to share our own beliefs, convincing friends to buy one disc rather than another, defending our own words and actions against critics, and self-interested statements of politicians and advertisers.

J focuses on beliefs. As a general rule, theories of J focus on the J of beliefs. This is in some way due to the influence of the seminal conceptualization of knowledge as *justified true belief*. If a belief is justified, there is something that justifies it, which is often defined as a *justifier*. If a belief is justified, then it has at least one justifier. Examples of justifiers for a belief are observational statements or other justified belief. For example, if a person is aware that her dog returned from the

yard smelling like garbage and that the trashcan outside had fallen over, the stink and the trashcan can be evidence for her belief that her dog went through the garbage. In that case, the justifiers are the person's awareness of the stink and the trashcan, and the belief that is justified is her belief that her dog went through the garbage.

Theory of J

Theory of J can be defined as the attempt to understand the J of statements and beliefs. There are various stances on what requires a *good* J. The core question for most of the theories of J is how to be sure or what we need in order to know or to assert that our beliefs correspond to the actual world. Differences among theories of J depend on what types of sources or evidence is required before a belief can be considered justified.

Theories of J are related to several epistemic features of beliefs. These include four crucial ideas for the theory of J: J, warrant, rationality, and probability. If A makes a claim *x* and B then casts doubt on it, A's next step would usually be to offer a J for *x*. Criteria that are used to consider a J as valid are where the boundaries are plotted between different theories of J and their philosophical foundations. Much of the debate in this field is focused on analyzing the nature of knowledge and how it relates to interlinked concepts, such as belief, truth, and J.

One negative theory of J is nonjustificationism. Nonjustificationism is the thesis that it is not possible to properly assign truth-values to sentences.

This view is often attributed to the work of Hans Reichenbach (1891–1953) and to the oeuvre of Karl R. Popper (1902–1994). Reichenbach thought that there was no logic of scientific method—that is, no proof or refutation. The basis for his claim was that there could be no refutation, as any theory could be protected from a putative refutation with some ad hoc maneuver.

In a different view from Reichenbach's conception, Popper also developed a theory on the growth of scientific knowledge *without* J. His conception is not justificationist, in the extent to which it does not properly claim to assign truth-values to propositions. All statements are hypothetical. Rather than asserting that a singular statement is truthful, Popper asserts that a basic sentence may be only provisionally accepted, and only if, this sentence is testable. Then, Popper introduces the following rule: consider only basic statements. Provisional acceptance of basic sentences—Popper claims—does not disqualify them as refutations of theories because for the most part we can agree on which basic statements we provisionally assume to be veridical. Secondly, Popper proposes the following rule: One should always substitute any theory which is contradicted by a basic statement by whichever new alternative theory that has a higher degree of falsifiability.

There are several different views on what entails a J. The first has to do with the classical opposition between an a priori J and an a posteriori J. A second discussion is on internalist versus externalist conceptions of a J. Finally, a third relevant concern is on the debate between foundationalism and coherentism in the theory of J.

A Priori J Versus A Posteriori J

There has been a long philosophical debate between empiricism and rationalism on what a *good* J is. Roughly speaking, rationalists defend an a priori conception of J, whereas empiricists defend an a posteriori conception of J. Each of these positions has been defended by a number of scholars. The oeuvre of Albert Casullo represents a skeptical vision on rationalism, whereas the work of Laurence Bonjour defends the rationalists' approach. On the other hand, other scholars

have attempted to challenge the cogency and significance of the a priori/a posteriori distinction, arguing that some warrants or justifications are neither a priori nor a posteriori.

On this concern, rationalism's argument is the following. Consider any belief or judgment affirming that *m* is inferentially justified on the basis of some other belief *n*. In order to be justified in believing *m* by inference from *n*, it is necessary (1) that one have good reasons for believing *n* and (2) that one have good reasons for thinking that *if n then m* (or *if n then probably m*). However, since experience might offer no direct support for the latter condition, any J for it must be (wholly or partly) a priori. Therefore, no experience can, on its own, directly justify a belief.

In opposition to the rationalist argument, empiricism responds by supporting the claim that one could justify general principles by induction from a number of experiences. Any J for it must be a posteriori. If successful, this would show that the empiricist is not vulnerable to the a priori problem. An example of this thesis can be seen through the empiricist J of mathematical beliefs. In this way, the inductivists' claim that appeals to an a priori source of J is *unnecessary* for the J of elementary mathematical propositions, in light of the available inductive support. Consider, for instance, the classical example of attempt to justify $1 + 1 = 2$. Confirming instances of this level of generality might involve counting one object and other one object, and then counting and finding them to be two in number.

Contrarily, the a priori conception of J affirms that those instances support the generality *only if* certain background assumptions are in place. In other words, mathematical propositions depend not just on *counting*: These propositions also depend on mathematical or logical claims to the effect that, for instance, some quantities are equal or unequal to others. Therefore, the J of some elementary mathematical claims will have to rely upon claims that are themselves not justified via induction. For example, the background assumption behind the sum according to which objects are stable can only be justified by assuming that the background assumption is already true. If a mathematical proposition must be justified by appeal to other a priori propositions, then this is no longer an

entirely a posteriori J—or a purely inductivist J—of elementary mathematical statements.

Therefore, a subject that has an a posteriori J, it essentially depends on statement of particulars or property instances. The belief that the abstract notion of circle exists, for example, might be justified by one's awareness of particular circles. Contrarily, a subject has an a priori J if one's J or evidence for a belief does not depend essentially on any particulars but only on universals. For example, an assertion about abstract logical relation of exclusion or incompatibility between the attribute of being circular and the attribute of being rectangular can offer a J for judging that no circles are rectangular, as well as J for the inference from the statement that something is a circle to its not being a rectangle.

Internalism, Externalism, and J

Internalism and externalism are two opposing ways of explaining the conception of an epistemic J throughout different dimensions. There has been a long debate in philosophy among these approaches. This dispute addresses one of the most fundamental questions in the discipline; that is to say, what is the basic nature of epistemic J. It is usually held that a valid epistemic status is not an attribute of a simple fact. Rather, a valid epistemic status is a condition of a belief *X if and only if* belief *X* is justified, as a rule, in virtue of some additional obtained condition.

Within the context of epistemology, internalism is the thesis asserting that everything necessary to provide J for a belief must be immediately available to an agent's consciousness. Oppositely, externalism asserts that the justificatory status of a belief can depend on other factors than those internal to the epistemic agent holding that belief. Internalism (or mentalism) affirms that what ultimately justifies any belief is some mental state of the believer. On the contrary, externalism asserts the thesis according to which there is something different to mental states that function as justifiers.

To put it succinctly, externalism holds that there are entities in the world which justify our beliefs or determine our meanings. Therefore, externalism is the theory that the contents of at least some of one's mental states are dependent in

part on their relationship to the external world. Conversely, internalism asserts that a belief can be explained by pointing to things which are internal to the believer's mind.

The central idea behind internalism is related to the nature of the justifier of a given justified belief. The justifiers for a given justified belief are those matters that constitute the person's J for that belief. They are those matters on which the person's current J is based, such as other beliefs, experiences, or states of affairs. For instance, consider a person telling us that a bear was in the neighborhood yesterday, and suppose that the testimony you receive from another person produces in you the justified belief that a bear was in the neighborhood. We can say that a *justifier* for this belief is the testimony one has received from the other person.

Internalists hold a rule according to which one has J for believing *S* on the basis of *S'* *only if* one has a J for believing that *S'* makes *S* probable. A repeated criticism that externalists have posed to internalists is on the cost of the vicious regress that internalism has.

Foundationalism, Coherentism, and J

Foundationalism has to do with those theories of knowledge resting upon *justified beliefs*. The two main theses of foundationalism assert that (1) *basic beliefs* exist, which are justified without reference to other beliefs and (2) *nonbasic beliefs* must ultimately be justified by basic beliefs. Contrarily, coherentism does not require a foundation: A belief can be justified if this belief is a member of a coherent set of beliefs.

There are two versions of foundationalism: a strong (or classical) version and a weak (or moderate) version. Strong foundationalism holds that there are basic beliefs that serve as foundations to anchor the rest of our beliefs. It advocates the infallibility of those basic beliefs and the use of deductive reasoning between beliefs. Weak foundationalism accepts provisional basic beliefs, and the use of inductive analysis between them. To put it succinctly, strong versions of this theory affirm that a justified belief is completely justified by basic beliefs, whereas weak versions of foundationalism maintain that justified beliefs can be justified by basic beliefs but can be additionally justified by other justified beliefs.

Strong foundationalism takes immediate perceptions and direct experience as good items or as acceptable foundations for basic beliefs. The strong version of foundationalism holds that beliefs about those matters do not need further support to be justified. Basic beliefs can be justified by some specific attribute of the belief itself, such as its being self-evident. Contrarily, weak foundationalism does not entail that a basic belief is *self-evident*. In a different way, the weak version of foundationalism asserts that it is reasonable to assume that basic beliefs are justified unless confirmation to the contrary is present.

On the other hand, foundationalism can be internalist or externalist. Internalist foundationalists hold that basic beliefs are justified by mental events. In other words, internalism holds that it is necessary for the means of J of a belief to be accessible to the believer. Contrarily, externalist foundationalists don't require that a believer's J for a belief must be accessible to that for it to be justified.

Coherentism opposes foundationalism, characterizing epistemic J as an attribute of a belief *only if* that belief is a member of a coherent set. As a theory of J, coherentism restricts justified belief to that which coheres with some specified set of beliefs. Coherentists assert that not all plausible underpinning J for a justified belief is supported finally for a basic belief. A core idea of coherentism as a whole is that J is a function of some connection between beliefs, none of which are privileged beliefs. On the other hand, different types of coherentism differ among themselves by the specific kind of relationship among beliefs such that every type is considered as coherence.

Therefore, varieties of coherentism emerge from the meaning of the concept of coherence (i.e., what it means for a system of statements to be coherent). A strong definition of coherence holds that a system of statements must include logical consistency. A moderate conceptualization of coherence requires some level of integration of the constituents of the system of beliefs. On the other hand, a theory *C* that uses only one explanation for divergent phenomena is more coherent than a theory *D* that explains those divergent phenomena using unrelated explanations. Additionally, a system *A* that uses only one explanation is more coherent than a system *B* that contains more than one unrelated explanation for the same phenomenon. Finally, to the extent to which the number of

phenomena explained by the system is greater, the greater will be its coherence.

Coherentism also opposes foundationalism's response to the so-called *regress argument*. The regress argument can be synthesized in the following way. Given some assertion *S*, it is plausible to require a J for *S*. If that J takes the form of another statement—let's call this statement *S'*—it is again feasible in logical terms to ask for a J for *S'*, and so on. Foundationalism's solution holds that the series of the questioning process finishes with certain statements having to be self-justifying. Coherentism's solution maintains that the series of questioning process forms a circle, so that each statement is ultimately involved in its own J. A third approach, infinitism's solution, maintains that the series is inherently infinitely long, with every proposition requiring a J by some other proposition, and so forth.

Foundationalism holds that it is unnecessary to require a J of certain propositions because they are self-justifying. Contrarily, coherentism asserts that it is always necessary to require a J for any proposition. From the coherentism's point of view, foundationalism proposes an unjustified place to end asking for J. Coherentism also maintains that foundationalism does not provide plausible reasons to think that certain beliefs do not require J. Foundationalism responds that it seems unreasonable to ask for justifications to certain statements (e.g., observational statements) such as, for instance, "it is raining."

Thus, coherentism rejects the reliability of the regression argument. The regression argument asserts that the J for a statement takes the form of another statement (e.g., *S''* justifies *S'*, which in turn justifies *S*), making J a linear process that ends at some statement that does not ask for J, that is, a justified belief. For coherentism, J for the belief in *S* is nonlinear: J is a *systemic* process. This means that *S''*, *S'*, and *S* are components of a coherent set of statements. *S''* and *S'* are not prior to *S* in epistemic terms. Rather, coherentism holds that *S''*, *S'*, and *S* work in an integrated manner within a set of statements to attain J.

J and the Theory of Argumentation

J is a core concept in the theory of argumentation. An argument can be defined—at least in a preliminary manner—as a set of related propositions.

Roughly speaking, in the context of the theory of argumentation, a distinction can be established between broad conceptualizations and narrow conceptualizations of a J. A broad conceptualization includes several kinds of defenses for an argument that seeks to prove that an argument is true. A narrow conceptualization defines J as a specific sort of reasoning facing an argument—that is, a reasoning that *is not* a way of arriving at ideas but rather a way of testing ideas critically. It is concerned less with how people think than with how they share their ideas and thoughts in situations that raise the question of whether those ideas are worth sharing. Thus, the question becomes the standards by which arguments succeed or fail to provide J for claims.

In terms of the *ends* orienting a J, there are two main types of J for an argument: strategic J and logical J. A strategic J of an argument pursues changing the belief of another person with relative or total independence regarding what is true—or perceived as true—by the speaker or writer. In turn, strategic J of an argument typically includes two different subtypes: (1) authoritative-based J of an argument and (2) emotional-based J of an argument. Contrarily to strategic J, logical J focuses on the plausibility of inferences and evidence to prove an argument. The logical J of an argument often uses syllogisms that include a major premise, a minor premise, and a conclusion based on the combination of the two premises.

Authoritative-based J of an argument appeals to the speaker's credibility, and establishing it depends less on their expertise and more on how they are depicted by themselves. To use credibility in favor of an argument, the speaker (or writer) must position themselves as an expert. Aristotle—the first to suggest an authoritative-based J of an argument that he called *ethos*—recommended that a speaker should use *powerful gesture, eye contact*, and so on “to position yourself as a leader.” Emotional-based J of an argument seeks to appeal to the emotions of the listeners, aiming to tempt them arouse their interest. An effective way of provoking passions is in appeal to values. Aristotle was also first to realize that language has a significant effect on emotion, and that certain words (such as fire, child, honor, or anger) can trigger senses and feelings. He called this kind of J for an argument, *pathos*.

Logical J of an argument—or *warrant*—is defined as a statement *authorizing the movement* from the *data* to the *claim*, using the terms of Stephen Toulmin. In his seminal approach to the theory of argumentation, a claim is a conclusion whose merit must be established. The classic Toulminian example of a claim is the following: If a person tries to convince a listener that he is a British citizen, the claim would be (1) “I am a British citizen.” Data or evidence is a fact one appeals to as a foundation for the claim. In that same example, the person introduced in (1) can support his claim with the supporting data (2) “I was born in Bermuda.” In order to move from the data established in (2), “I was born in Bermuda,” to the claim in (1), “I am a British citizen,” the person must supply a warrant to bridge the gap between (1) and (2) with the statement (3) “A man born in Bermuda will legally be a British citizen.” These aspects will be considered as a good argument *if and only if* they can succeed in providing good J to a claim, which will stand up to criticism and earn a favorable verdict.

The Toulminian model of argumentation includes three other related concepts: *backing*, *rebuttal*, and *qualifier*. A backing must be introduced when the warrant itself is not convincing enough to the readers or the listeners. In the seminal example given above, if the listener does not deem the warrant—or J—in (3) as credible, the speaker will supply the legal provisions as backing statement to show that it is true that “A man born in Bermuda will legally be a British citizen.” A rebuttal is a statement recognizing the restrictions to which the claim may legitimately be applied. The rebuttal, in the aforementioned example, is exemplified as follows: (4) “A man born in Bermuda will legally be a British citizen, unless he has betrayed Britain and has become a spy of another country.” Finally, a qualifier is a word or phrase expressing the speaker's degree of force or certainty concerning the claim. Such *words* or phrases include “probably,” “possible,” “impossible,” “certainly,” “presumably,” “as far as the evidence goes,” or “necessarily.” The claim (5) “I am definitely a British citizen” has a greater degree of force than the claim (6) “I am a British citizen, presumably.”

The work of Carl R. Rogers and his disciples introduces what might be considered a third type of J for an argument: a deliberative-based J. This type of J can be separated into five components: (1) an

introduction to the problem and a demonstration that the opponent's position is understood; (2) a statement of the contexts in which the opponent's position may be valid; (3) a statement of the speaker's position, including the contexts in which it is valid; (4) a statement of how the opponent's position would benefit if he were to adopt elements of the speaker's position; and (5) if the speaker can show that the positions complement each other, that each supplies what the other lacks, so much the better.

Rodolfo Sarsfield

See also Consistency; Falsifiability; Logical Theory; Refutation; Relative Consistency; Soundness

Further Readings

- Aristotle, R. (2014). The Internet classics archives (W. R. Roberts, Trans.). <http://classics.mit.edu/Aristotle/rhetoric.html> [350 B.C.E.].
- BonJour, L. (1980). Externalist theories of empirical knowledge. *Midwest Studies in Philosophy*, 5, 53–73.
- BonJour, L. & Sosa, E. (2003). *Epistemic justification: Internalism vs. externalism, foundations vs. virtues*. Blackwell.
- Bovens, L., & Hartmann, S. (2003). *Bayesian epistemology*. Oxford University Press.
- Casullo, A. (2003). *A priori justification*. Oxford University Press.
- Casullo, A. (2012). *Essays on a priori knowledge and justification*. Oxford University Press.
- Goldman, A. (1979). What is Justified Belief? In G. S. Pappas (Ed.), *Justification and knowledge*. D. Reidel.
- Hasan, A. (2014, January). Review of Albert Casullo, *Essays on a priori knowledge and justification*. *Notre Dame Philosophical Reviews*.
- Hawthorne, J. (2007). A priority and externalism. In S. Goldberg (Ed.), *Internalism and externalism in semantics and epistemology*. Oxford University Press.
- Madison, B. J. C. (2010). Epistemic internalism. *Philosophy Compass*, 5 (10), 840–853.
- Olsson, E. J. (2005). *Against coherence*. Oxford University Press.
- Pappas, G. (2014). Internalist vs. externalist conceptions of epistemic justification. In E. N. Zalta (Ed.), *The Stanford encyclopedia of philosophy*. <http://plato.stanford.edu/archives/fall2014/entries/justep-intext/>
- Popper, K. R. (1959/1935). *The logic of scientific discovery*. Heinemann.
- Reichenbach, H. (1938). *Experience and prediction: An analysis of the foundations and the structure of knowledge*. University of Chicago Press.
- Rogers, C. R. (1961). Communication: Its blocking and its facilitation. In C. Rogers (Ed.), *On becoming a person: A therapist's view of psychotherapy*. Constable.
- Russell, B. (1912). *The problems of philosophy*. Oxford University Press.
- Toulmin, S. (2003/1958). *The uses of argument*. Cambridge University Press.

K

KINETIC (MOLECULAR) THEORY

The kinetic theory of an ideal gas establishes a connection between mechanics and thermodynamics, suggesting that thermodynamic quantities can be understood in terms of mechanical properties averaged over a population of particles. An ideal gas is one of the simplest models for a physical system consisting of multiple particles and is an excellent approximation to many real gases. Moreover, as this entry shows, even this very simple model introduces some long-standing conceptual issues in statistical physics, such as the nature of assumptions concerning randomness and the need for a discrete set of states a system can occupy.

Connecting Mechanics and Thermodynamics

In the early 19th century, the basic Newtonian mechanics of objects was well understood and could plausibly have been thought to be the fundamental physics underlying how the universe worked. The industrial revolution, however, prompted the invention of thermodynamics to handle systems considerably more complex than those captured by simple mechanical models. The working fluid of a steam engine can be described in terms of variables such as its pressure P , volume V , absolute temperature T , and entropy S . Some of these variables can be straightforwardly connected to basic mechanics: $P = FL/A$, or the

pressure is the perpendicular component of a force uniformly distributed over an area A . Steam somehow pushes against the walls of its container. But other variables, such as temperature, had no such translation into mechanics.

Consider the equation of state of an ideal gas, $PV = nRT$, where R is a constant and n , the number of moles, is a measure of the amount of gas. This equation combines variables that can be linked to mechanics and those that do not have an obvious connection. Now, thermodynamics can certainly be formulated as a stand-alone theory, without any grounding in basic mechanics. However, relationships like the equation of state suggest that there might be a connection. Indeed, physicists can seek to *reduce* thermodynamics to a more fundamental mechanical theory, providing a more unified description of nature and a deeper understanding of why thermodynamics works as it does.

Physicists perform this reduction through *statistical mechanics*: Thermodynamics becomes a description of systems with very large numbers of components. Thermodynamic variables are rooted in the average properties of a population of components. They are statistical variables.

The ideal gas equation of state might then be understood as describing the population of an immense number of particles. The volume V , for example, would then be the volume occupied by the whole population of particles, rather than the volume of any single particle. The pressure P would be due to lots of particles constantly bouncing off the walls of the container, and n would have to do

with the number of particles—all statistical, population-level descriptions. This leaves temperature to be figured out. In any case, the very basic statistical mechanics represented by the kinetic theory of an ideal gas requires an atomic theory of matter to be correct. Instead of being an infinitely divisible continuum, the ideal gas has to consist of a discrete, countable number of particles. The microscopic behavior of each individual particle will be captured by basic mechanics, and the statistical behavior of the population as a whole should reproduce the results of a thermodynamic analysis.

The ideal gas is one of the simplest possible models of a multi-particle system. Gas molecules very weakly interact with one another; in an ideal gas model, the particles do not interact with each other at all, undergoing completely elastic collisions with the walls of the container. With these assumptions, the kinetic theory calculation is very straightforward.

Say the container is a cube with each side having length L . Focusing on a single particle bouncing off only one of the six walls, each collision will result in a momentum change of magnitude $\Delta p_x = mv_x - (-mv_x) = 2mv_x$, where m is the mass of the particle and v_x is the x -component of its velocity. The particle will then travel to the opposite wall, bounce off that, and make its way back for another collision. The time interval between each collision is then $\Delta t = 2L/v_x$. Since force is the rate of change of momentum, the time-average of the force delivered to this one wall by one particle must be $F_x = \Delta p_x / \Delta t = mv_x^2 / L$.

Now, the total force will be $\Sigma F_x = Nm\overline{v_x^2} / L$, where N is the number of particles and $\overline{v_x^2}$ is the average of the square of the x -component of the velocities of the particles. The pressure is this force divided by the area of one side, L^2 . Therefore $P = \Sigma F_x / L^2 = Nm\overline{v_x^2} / L^2$. The volume of the container is L^3 , and $Nm\overline{v_x^2}$ looks like it is related to the total kinetic energy, $K = 1/2 Nm (\overline{v_x^2} + \overline{v_y^2} + \overline{v_z^2})$. But there is no distinction between the x , y , and z directions, and therefore $\overline{v_x^2} = \overline{v_y^2} = \overline{v_z^2}$. Putting all this together, the result is $PV = 2/3K$.

The ideal gas equation of state, established through empirical means with actual gases, is

$PV = NkT$, now phrased in terms of the number of particles and the Boltzmann constant k , such that $nR = Nk$. All this means that in an ideal gas, temperature must be related to the total kinetic energy of the population of particles so that $K = 3/2NkT$. Alternatively, the temperature of an ideal gas tells us that the average kinetic energy of the particles is $\overline{K} = K/N = 3/2kT$.

This result of kinetic theory, linking microscopic mechanics with macroscopic thermodynamics, applies to the ideal gas only. Physicists would later establish a more general definition of temperature in terms of how entropy changes as the total energy U of a system increases; $1/T = \partial S / \partial U$. In statistical mechanics, entropy is related to the number of microscopic states that produces the same set of macroscopic thermodynamic variables that define a system. For an ideal gas, however, the only energy in the system is kinetic energy, and it is possible to establish the mechanical meaning of temperature without dealing with the complexities of entropy.

The kinetic theory of an ideal gas, therefore, is just a first step in reducing thermodynamics to statistical mechanics. The connection kinetic theory establishes between temperature and kinetic energy in a gas, however, is highly significant: with a very simple, exactly solvable model, it shows that a connection can be made. Indeed, the stripped-down, simplified nature of an ideal gas model helps emphasize the fundamental features of a statistical approach to large many-particle systems. Complications can be built in later. Moreover, an ideal gas is an excellent approximation to most gases under everyday conditions, so much that engineers typically take the gases they work with to be ideal gases under most circumstances.

Conceptual Issues

While kinetic theory can be thought of as an early, and still useful, application of statistical mechanics, the vastly simplified model of an ideal gas also leads into some of the conceptual issues that a more mature statistical physics has to confront. The most prominent of these issues are the nature of the randomness assumptions in statistical models, the meaning of probabilities in the models, and the necessity of discrete state spaces for well-defined concepts.

Randomness and Probability in the Ideal Gas

Since basic kinetic theory refers to averages in a population of particles, questions about how the positions and velocities in an ideal gas are distributed arise. In a more developed statistical mechanics, the velocities have the well-known Maxwell–Boltzmann distribution. The positions of the particles are random, uniformly distributed throughout the volume. But all this is not part of kinetic theory. Indeed, there is a question of why randomness should appear in the initial conditions.

One possible answer is that if all initial positions for the ideal gas particles are equally likely, the probability for anything but a random initial distribution is exactly zero. If, for example, the particles in an ideal gas started arranged on a cubic lattice, with identical velocities, such a configuration might produce odd results such as unequal pressures on the container walls. However, not only this, but any ordered arrangement has zero probability.

Statistical physics then proceeds beyond kinetic theory by introducing the fundamental assumption that all microscopic states of the population of particles, constrained by appropriate requirements such as equal total energies, must be equally probable. Such an assumption can be hard to justify beyond the fact that it works, since the Newtonian mechanics that is supposed to underlie thermodynamics is completely deterministic.

The probabilities that are fundamental to statistical physics raise similar questions. For example, entropy can be defined through the information entropy as $S = -k \sum p_i \ln p_i$, which refers to the probabilities p_i of physical states. Indeed, so defined, entropy measures a lack of information about the state of a system. But there is something peculiar about such a perspective. After all, the thermodynamic definition of entropy has $\Delta S = dQ/T$, a quantity related to heat transfers that can be measured in a lab through calorimetry. Does entropy say something about the lack of knowledge of a physicist—are the p_i epistemic probabilities? Or are the probabilities, like the thermodynamic entropy, objective features of physical systems, independent of the physicist?

One approach to supporting the fundamental assumption is ergodic theory, which studies whether the states occupied by a population of particles evolving in time are a representative sample of all the states assumed to be equally probable. Many dynamical systems, although deterministic, are sensitive to initial conditions in a manner that means that any uncertainty in the knowledge of the state of a system grows exponentially with time. In effect, the systems become randomized. If ideal gas particles are allowed to collide with each other or if the walls of the container have irregularities, the result will be a similar degradation of knowledge about the precise state of the system—there will be no alternative but to describe the gas by using macroscopic, statistical variables.

Processes that lead to loss of information can bridge the gap between deterministic mechanics and assumptions of randomness, and the gap between epistemic and objective probabilities. Nevertheless, not all details are clear, and such conceptual issues have continued to interest both statistical physicists and philosophers of science.

The Need for Discrete State Spaces

Kinetic theory connects thermodynamics and particle mechanics by discretizing the ideal gas: taking the gas to be a countable collection of particles. Another conceptual problem that arises at this stage comes from the fact that the process of discretization is incomplete: The classical ideal gas still has a continuous, nondiscrete state space. This leads to infinities and other difficulties that require quantized states for their resolution, hinting at a deeper connection between statistical and quantum physics.

The entropy of a system is defined in terms of the information entropy or, equivalently, in terms of the number of microscopic states corresponding to macroscopic thermodynamic states. The states of the classical particles that make up an ideal gas are defined by the particle position and momentum. But both the position and momentum are continuous variables. Therefore, there are an uncountable number of states of an ideal gas, and defining an entropy through state-counting becomes impossible.

In the late 19th and early 20th centuries, thermodynamic theories ran into some notorious problems with infinities, such as the “ultraviolet catastrophe” and the divergence of specific heats due to the equipartition theorem. In hindsight, these were evidence of the need for quantum mechanics. Early workarounds for these problems recognized that discretizing the state space led to finite results that could fit observations.

An early attempt to discretize the state space of an ideal gas produced the Sackur–Tetrode equation for the ideal gas entropy, which works very well at temperatures where the ideal gas particles behave classically. The discretization is, however, somewhat crude, and at temperatures close to zero, the Sackur–Tetrode entropy becomes negative, which is absurd.

A more proper quantum treatment of an ideal gas uses the wave functions for particles confined to a box. While the wave functions and the energy levels in this case are exact solutions, obtaining most quantities of interest in statistical mechanics requires the partition function, which is not exactly summable. The quantum ideal gas might be the modern equivalent of kinetic theory, but it relies on high-temperature approximations and numerical calculations to produce results. Unlike the easily visualizable, mathematically tractable model of the kinetic theory of a classical ideal gas, even the simplest possible quantum multiparticle system analogous to an ideal gas requires more advanced techniques and is not very intuitive.

Taner Edis

See also Information Theory; Physics, 19th Century; Physics, Quantum Theory; Physics, Thermodynamics

Further Readings

- Gaspard, P. (1992). Diffusion, effusion, and chaotic scattering: An exactly solvable Liouvillian dynamics. *Journal of Statistical Physics*, 68, 673–747.
- Rosenkrantz, R. D. (Ed.). (1989). *E. T. Jaynes: Papers on probability, statistics and statistical physics*. Kluwer Academic.
- Schroeder, D. V. (2000). *An introduction to thermal physics*. Addison-Wesley.

KNOWLEDGE

Theories of knowledge serve as a foundation logic for schools of thought about human experience and the world. Schools of thought can be distinguished by their varied beliefs about how knowledge is defined, how it is acquired, the differences among cognitive states involved, and how it is represented. There is a great benefit to reducing vast amounts of discrete information, or experience, to form broad generalizations that help filter the continuous barrage of new perceptions. Persons unable to impose such filters on their experience become overwhelmed, as seen in some autistic people who panic and adopt insulating behaviors to limit the scale and intensity of stimulation and the ambiguity of sensation and experience. Thus there is a high value placed on knowledge that is consistent within a cohesive framework, resulting in worldviews that can allow new information to be assimilated and can expose fallacy while maintaining a sense of order. This entry discusses the nature of knowledge generally, including the distinction between knowledge and belief. It provides a brief overview of western historical notions of what constitutes knowledge, from ancient Greece through the medieval period, Cartesian mind–body dualism, the scientific revolution, the development of positivism and modernity. The entry concludes with a description of the more recent contributions of cognitive theory, developmental psychology, and other fields, and the connections of knowledge theory with other areas of investigation such as ethics, as well as information theory in the age of Big Data.

Defining Knowledge

In common usage, *knowledge* is an unambiguous term that indicates a framework for understanding information. Whereas information may be regarded as passive data, once it is understood in meaningful context, it contributes to a knowledge base. That knowledge base can then be applied to problems, in which case it becomes a tool for critical thinking. A knowledge base is foundational,

but it is not considered finite because doubts and curiosity persist to inspire further investigation. This is implied when an argument is qualified by the speaker's admission that it is "to the best of my knowledge." Thus the root of the word reveals its core attribute: the activity of knowing. Knowledge is thus assumed to be within the scope of human endeavors. It is further regarded as a continuous effort as people find order and meaning in their existence. The accusation that "you don't know what you are talking about" suggests not knowing enough information or not integrating it logically. That may be as nuanced a description of knowledge as most people care to use, but a much closer examination of knowledge by its nature and role in human experience is a primary focus of philosophical inquiry, broadly termed *epistemology*, variously defined as theories the nature of knowledge or the nature of knowing (see the entry on Understanding, this encyclopedia).

Knowledge is rendered ambiguous due to its several dimensions as revealed by responses to the following questions: What is the source of the knowledge—for example, innate, divine inspiration, authority, social convention, deductive reasoning, scientific observation? How robust is the knowledge? Is it absolute and eternal or relative and fluid? How is the knowledge organized, by isolated propositions or integrated frameworks? *Epistemic cognition* refers to the understanding one has about the nature of knowledge and learning. A causal chain typically considers the quality of information, the neutrality of reasoning, the degree of honest intention, and the influence of other factors, resulting in a familiar list of virtues associated with intellectual integrity: honesty, rationality, reliability, trustworthiness, and transparency.

The Nature of Knowledge

The Greek base word *episteme* (γνῶσις) means *to stand up*, for knowledge was assumed to be stable and therefore reliable no matter how it was acquired because the world to be known was assumed to be stable. Classical education focused on perennial laws and principles. The certainty with which these concepts were regarded as

universal corresponded to the authority granted to the source of this knowledge. The classical definition of knowledge requires three conditions: It must be believed, it must be true, and it must be justified, that is, *justified true belief*. Beliefs, or *doxo* (δόξα), are regarded as opinion, not fact, and are therefore changeable and subject to influence. In contrast, knowledge is expected to stand up to distortion and challenge. Justification constitutes the rational explanation and evidence that establishes the statement of knowledge, thereby providing an argument for its stability. The justification distinguishes knowledge from understanding; for example, Galileo's logical conclusions concerning astronomy, based on observation, were not accepted as knowledge until Johannes Kepler provided a supporting theory and mathematical explanation. By contrast, naive or intuitive understanding lacks explicit explanation, although there is lively debate on this point among philosophers. Some insist that knowledge can be a collection of separate facts, whereas understanding requires incorporation of knowledge into a cohesive framework.

The syllogism is the tool of classical analysis, examining phenomena in terms of the necessary and sufficient conditions specified in an authoritative definition of the concept. Implied here is that it is useful to define essential characteristics in order to make sense of the world, and further implied is that philosophical inquiry serves a broader purpose of trying to predict what may occur or explain why something has occurred. This view is called *positivist* because it assumes that one can be positive about the reality one hopes to know. In contrast, postmodern theories argue that the nature of the world is complex and changing and therefore knowledge cannot be any more stable than the world it describes. Here the emphasis is more existential; that is, that knowledge is a conceit shared among people who have found similar meanings in their experiences. A postmodern departure from the assumption that universal truths can be identified and then used to make sense of novel situations is found in critical theorists such as Michel Foucault (1926–1984), who argued that no truth exists outside of the individual's context and cultural

interpretation, thus knowledge cannot be agreed upon, let alone generalized, because of the unique combinations of experience.

Acquiring Knowledge

The stability of reality is a different issue from the stability of knowledge. In ancient Greece, Plato's idealism and Aristotle's realism diverged regarding the nature of the world and the ways it may be understood, contrasting the mind with the body for apprehending truth. The *sensible*, which can be perceived directly from immediate, concrete, sensory experience, is qualitatively different from the *intelligible*, which can be understood indirectly through abstract thought. Plato dismissed sensation, which he thought could only result in opinion, not knowledge. Socrates, his teacher, sought to discover the essence of a concept, not a catalogue of its examples nor even the variations in context that would cause doubt. In Plato's Socratic dialogue the *Meno*, the student becomes knowledgeable by touring his existing, or foundational, understanding and drawing rational conclusions about the only dimension that existed for his teacher: abstract ideas that do not rely on a physical manifestation. The abstract concept was regarded as the knowledge base. The legacy of Plato can be seen in modern thinkers, beginning with René Descartes' (1596–1650) skepticism, Carl Gustav Jung's (1875–1961) theory of archetypes, and, more recently, the work of linguist Noam Chomsky, all of whom argued in various ways that certain concepts are innate, that is, a priori. Immanuel Kant (1724–1804) had introduced a transcendental idealist argument in hopes of resolving the mind–body issue, but his solution has not been universally accepted and the debate continues over the nature of mind and the relative position of the mental and material domains in the development of knowledge.

A large part of Aristotle's legacy is *empiricism*: the scientific tradition of perceiving multiple instances and reducing the data to a central tendency in order to propose or test a theory about some property of reality. Instead of assuming inherent universal concepts that are innately known by all people, this approach requires experience in order to discover essential characteristics.

The English philosopher John Locke (1632–1704) concurred, arguing further that we are not born with innate concepts but rather that the mind is a *tabula rasa*, or blank slate awaiting personal experience. Differences among the types of knowledge to be acquired are significant to education; that is, society's intentional efforts to expand the competence of the untutored. In 1859, the biologist and philosopher Herbert Spencer (1820–1903) argued for science to be included in the public school curriculum in his influential essay "What Knowledge Is of Most Worth." He challenged the traditional focus of education on classical analyses of universal concepts memorized by rote in favor of self-evident truths with relevance to students' personal lives. John Dewey (1859–1952), one of the most influential of American philosophers and educators, objected to false dichotomies that would mutually isolate functions of the mind and body. The educational implication was that knowledge could not simply be transmitted in the traditional way of lecture and memorization, but that authentic experience in a meaningful context is required in order to confront a misconception or lack of knowledge and thereby construct a new understanding. Folk wisdom accepts intuitive insight as a source of knowledge.

An interactive theory of knowledge is found in developmental psychology, as in the pioneering psychoanalyst Sigmund Freud's (1856–1939) model that describes not just an innate and predictable sequence of development but distortions of understanding resulting in abnormal behavior. Swiss developmental psychologist Jean Piaget (1896–1980) observed that children begin with concrete interaction with physical objects, or interobjectivity, and develop gradually more abstract reasoning. Lev Vygotsky (1896–1934), an influential Russian psychologist, emphasized intersubjectivity, that is, interaction with people, and current thinking respects the influence of context and self-image on an individual's capacity and tendencies to learn. In the parlance of cognitive psychology, a distinction is made between automatic versus controlled thought processes which have been associated with different neurological locations. The executive functions, controlled by the prefrontal cortex, include tools for knowing: vocabulary and working memory.

Representing Knowledge

Current tradition among educators divides experience into three realms: cognitive (thinking), affective (feeling), and physical (doing). These categories are useful for analyzing multiple dimensions of the same experience, resulting in different sets of knowledge. *Ontology* is the study of classifying knowledge. The medieval theologian Thomas Aquinas (1225–1274) sorted all knowledge on the basis of its need for matter and motion in order to be known. The American psychologist Howard Gardner has suggested multiple ways of knowing, or Multiple Intelligences, with which to experience the world and to represent one's knowledge of the world. This theory posits that any concept can be understood and represented in multiple ways, thus supporting education reforms that insist all children can learn if the instruction can be appropriately differentiated and delivered in an engaging way. This extends Jerome Bruner's (1915–2016) hypothesis that any knowledge can be learned if reduced to a developmentally appropriate conceptual unit, but it begs the question of how to determine whether the academic objective was achieved.

Interpreting evidence of learning is in the realm of educational assessment, which highlights the complexity of knowledge. The concern for valid assessment prompted a group of evaluators to develop a taxonomy of cognition, popularly named for the leader of the team, Benjamin Bloom (1913–1999). The original *Bloom's Taxonomy* identified the lowest level of cognition as *knowledge*, meaning simple facts with the next level as broader categories for grouping the knowledge, or *comprehension*. These two levels plus an *application*, which can transfer a concept to a new situation, constituted the lower levels of thinking and the foundational knowledge base necessary for the higher levels of thinking: *Analysis* involves sorting aspects of a complex body into different components and understanding how they work together; *synthesis* involves connecting concepts from different entities in order to create something new; *evaluation* involves making decisions and exercising judgment. Half a century later, a revision was published that refined the taxonomy, first of all by renaming the levels in language suggesting

the active function of thinking: (1) recalling, (2) explaining, (3) applying, (4) analyzing, (5) judging, and (6) creating. More significantly, the term *knowledge* was changed from designating the lowest level of cognition, into four dimensions that infuse all levels of thinking: factual, conceptual, procedural, and metacognitive. These are ontological divisions, that is, ways to sort knowledge according to its own characteristics; each of them is further articulated:

Factual knowledge includes knowledge of terminology (currently called *academic language* according to teacher certification standards), and knowledge of specific details and elements. This is the easiest category of knowledge to gather and to assess, for it tends to have *one right answer*, and is thus often tested with selected response items (e.g., multiple choice; true or false) which involve the lowest level of cognition, recall. *Conceptual knowledge* includes knowledge classifications and categories, or principles and generalizations; and theories, models, and structures that can thus group factual details by concept. Constructed, or extended, response items (e.g., essays, diagrams) can demonstrate conceptual understanding, typically requiring facts as illustration and deductive reasoning for justification. Mathematics and sensory perception offer the greatest certainty, which diminishes as the source of information recedes from matter and motion. Memory and that which cannot be observed directly are less certain. Abstract, philosophical issues are the most dubious as well as the most difficult to explain but are nonetheless valued for providing a stable worldview regarding one's purpose or, for many, the existence of a supreme being without the burden of scientific proof. Factual and conceptual knowledge are considered propositional, that is, they are statements of truth about the world beyond the individual's personal or procedural knowledge.

Procedural knowledge involves the practical application of factual and conceptual knowledge, using subject-specific skills and procedures, methods, and techniques as well as knowledge of criteria for deciding when to use them. Performance assessments, or demonstrations, are needed to assess procedural knowledge; mastery of attendant facts and concepts cannot be mistaken for procedural competence. For instance, knowledge of the

vocabulary specific to cooking and knowledge of the concepts of heat and seasoning are necessary but insufficient for knowledge of how to make dinner.

The final dimension, *metacognitive knowledge*, includes knowledge of strategies for learning and of cognitive tasks as well as knowledge of one's own self in terms of strengths and weaknesses in understanding. As the ancient philosopher Laozi (Lao Tsu) said: Knowledge of others is intelligence; knowledge of oneself is true wisdom. Metacognition serves the purpose of critical rationalism by reserving doubt for any pronouncement and by acknowledging a need for further inquiry. While self-knowledge was valued in order to "Know thyself," and "an unexamined life is not worth living," John Locke suggested that "internal sense" was not just a way to process external sensation, but a source of original ideas. Cognitive psychologies focus on the importance of self-awareness and their therapies center on dissonance between belief and experience, that is, knowledge. Canadian American psychologist Albert Bandura (1925–) contributed the concept of *self-efficacy*; that is, confidence in one's capacity to succeed. Another predictor of success is *epistemic cognition*; that is, understanding the nature of knowledge itself. This includes epistemic aims, such as the degree to which one wants to understand something, and epistemic values. The primary tool of metacognition is reflection, the deliberate exercise of self-awareness, valued highly by Socrates, who reputedly said "An unexamined life is not worth living." Reflection can be seen in *cybernetics*, a model for thinking systemically that focuses on feedback loops as the deliberate invitation of information that may not conform to one's assumptions of what is. This includes a political dimension, for consulting stakeholders is regarded as essential for effective implementation of any reforms. Metacognition can be seen in community-wide awareness of the variance of perspective within its collective consciousness.

Knowledge in Context

Specialized fields of knowledge have emerged, such as psychology, education, and law, each of which has its own philosophical traditions within its scope of interest. *Knowledge management* is

the study of information technology and the role of people who must store and analyze data. Peter Drucker (1909–2005), a business analyst, who coined the term *knowledge worker*, noting the change in economic context from farm to factory and now to information. *Truth in advertising* focuses on the expectation that any information provided be verifiably true; ambiguity and omissions are considered as reprehensible as falsification. Journalistic integrity prohibits plagiarism, for which careers are terminated. Intellectual property rights continue to be debated in the context of democratic *need to know* and scientific verification requirements.

Ethical concerns pervade the topic of knowledge, both its presentation and access to it. First-world citizens democratically demand that they have a need and a right to know, and there are common standards regarding the selective, distorted, or even deliberate misplacement of that which is true. Horace Mann (1796–1859), pioneer of professional teaching, called education "the great equalizer," a nod to the old saying that "Knowledge is power." Brazilian reformer Paolo Freire (1921–1997) proposed a *pedagogy of the oppressed* to challenge traditional exclusion of impoverished children from active engagement in learning. Academic standards and how they may be measured are major political issues because most formal schooling is funded publicly and accounts for a large part of state budgets; school reform may be presented as consumer protection legislation, thus equating knowledge with a commodity. Current educational reform has promoted the adoption of standards to define the knowledge base desired upon graduation, with Common Core standards of literacy and numeracy and Next Generation Science Standards gathering national momentum.

The recent expansion of computing capacity has resulted in massive amounts of data that can be analyzed with increasingly sophisticated statistical software, leading to a phenomenon undermining the traditional scientific model of testing hypotheses and generating theory. These theories are used to summarize principles thought to help explain phenomena in the real world, guiding future actions to influence that world. However, search engines such as used by Google now employ algorithms that identify ads compatible

with the user, having no conception of the meaning in those ads. This can render the usefulness of a theory moot because the validity of the data analysis inspires belief in the findings even without arguing philosophical merit. Thus knowledge for narrow epistemic aims may be justified without meaningfulness to other aims, such as reliable propositions of truth that can be generalized to other contexts. *Gettier cases* are scenarios contrived to illustrate unjustified true beliefs.

Naomi Jeffery Petersen

See also Concept; Epistemology; Justification; Truth; Understanding

Further Readings

- Anderson, L., & Krathwohl, D. (Eds.). (2001). *A taxonomy for learning, teaching and assessing: A revision of Bloom's taxonomy of educational objectives*. Pearson Longman.
- Chisholm, R. (1989). *Theory of knowledge*. Prentice Hall.
- Day, S., & Goldstone, R. (2012). The import of knowledge export: Connecting findings and theories of transfer of learning. *Educational Psychologist*, 47(3), 153–176.
- Drucker, P. (1993). *The post-capitalist society*. HarperCollins.
- Gelman, S. (2011). Concepts and folk theories. *Annual Review of Anthropology*, 40, 379–438. <https://doi.org/10.1146/annurev-anthro-081309-145822>
- Hofer, B., & Pintrich, P. (1997). The development of epistemological theories: Beliefs about knowledge and knowing and their relation to learning. *Review of Educational Research*, 67(3), 88–140.
- Kuhn, D., & Park, S.-H. (2005). Epistemological understanding and the development of intellectual values. *International Journal of Educational Research*, 43, 111–124.
- Schommer-Aikins, M., & Spender, J. C. (2004). Knowledge fields: Some post-9/11 thoughts about the knowledge-based theory of the firm. In C. Holsapple (Ed.), *Handbook on knowledge management 1: Knowledge matters*. Springer.



LAWS, SCIENTIFIC

Scientific laws are central to science, but what *is* a scientific law? It is quite easy to give some examples: Coulomb’s law of electrostatic attraction, Snell’s law of refraction, the Schrödinger equation for the evolution of a quantum-mechanical system. But what do these things have in common in virtue of which they get to count as scientific laws? What *is* a scientific law—in other words, an approximation of a law of nature? While this is primarily a philosophical question, it should be no less interesting to scientists and students of science than, say, the foundations of quantum mechanics. Of course, there are many scientists to whom the latter is of no interest: the “shut up and calculate” types. But if one is not content with merely using science as a tool and would like to go further and investigate what science says the world is really like, then one should find something of interest in our philosophical question: “What is a scientific law?”

Simple Regularity View

To get a sense of why the question “what is a scientific law?” is puzzling, it will help to consider a very simple answer first. According to the *simple*

Acknowledgment: This work has received funding from the European Research Council (ERC) under the European Union’s Horizon 2020 research and innovation program as part of the MetaScience project, Grant Agreement No. 771509.

regularity view, laws of nature, or scientific laws, are true generalizations, roughly of the form “all Fs are Gs,” in other words *regularities* in nature. A generalization of the form “all Fs are Gs” is *true* if, and only if, there are no Fs that are not Gs.

The following are thus some candidates for lawhood:

List 1

- All swans are white
- All marsupials have a pouch
- All philosophers like coffee

There are black swans in Australia, the short-tailed opossum is a marsupial that does not have a pouch, and (apparently) some philosophers do not like coffee. So, none of the examples in List 1 are laws according to the simple regularity view because they have exceptions and hence are not true.

Now consider:

List 2

- Everything travels at or below light speed
- All massive objects exert a gravitational force on all other massive objects
- All uranium spheres are less than 1 mile in diameter

The examples in List 2 are true generalizations; they hold at all times and places in the universe without exception. So, according to the simple

regularity view, all the examples on List 2 are laws of science.

So far, so good for simple regularity view; it provides a simple answer to our question “What is a law of science?”—namely, to be a law is to be a true generalization, and it seems to appropriately classify our examples; plausibly, it is not a law that all philosophers like coffee but it is a law that everything travels at or below light speed.

But now consider the following: “All gold spheres are less than 1 mile in diameter.” It is certainly the case that no gold sphere greater than 1 mile in diameter has ever been discovered. Let’s assume, then, that there really are no gold spheres greater than 1 mile in diameter (there never have been and never will be). That is to say, “All gold spheres are less than 1 mile in diameter” holds at all times and places throughout the universe; it is a *true generalization*. According to the simple regularity view, it is then a *law of nature* that all gold spheres are less than 1 mile in diameter.

Although we may be happy to admit that it is a law that all *uranium* spheres are less than 1 mile in diameter, we probably do not want to say that it is a law that all *gold* spheres are less than 1 mile in diameter. The former has to do with the fact that uranium is an unstable *radioactive* element, whereas the latter seems like a mere accident.

The simple regularity view ultimately fails as an account of laws because it is unable to distinguish accidentally true generalizations from lawful ones. There must be more to being a scientific law than simply being a true generalization, but what more is needed?

Sophisticated Regularity View

According to a more sophisticated version of the simple regularity view, laws are true generalizations, but not just any true generalizations count as laws. Rather, laws are those true generalizations that make up a particularly efficient systematic description of the entire universe. Let’s explore this idea in a little more detail.

Anyone can begin to describe the universe by just listing things that she knows: Grass is green, the Eiffel Tower is in Paris, electrons are negatively charged, and so on. If we knew absolutely everything there is to know about the entire universe, past, present, and future, we could continue

this list until we had described the universe in its entirety.

There are more and less efficient ways of describing the universe. Including “grass is green” in our descriptive list is more efficient than specifying each blade of grass by some coordinate system and saying of it that it is green, but is perhaps less efficient than saying that all chlorophyll-containing plants are green. The most efficient description of the universe is the one that is the *strongest*, by which we mean that it is highly informative because it logically implies the most information about matters of fact (such as that grass is green, electrons are negatively charged, etc.) in the *simplest* manner possible; that is, by including as few basic, nonderived, statements possible.

Strength and *simplicity* are what we might call *virtues* of a description of the universe because a description is better the more informative it is and the simpler it is. But these virtues *compete*. The stronger an overall description of the universe is, the less simple that description will be because it will contain a greater number of basic statements. And the simpler a description is—in other words, the fewer basic statements it contains—the weaker it will be because it will have fewer resources to directly say, or logically entail, true things about the universe.

The best description of the universe will be the one that strikes the optimal balance between strength and simplicity. According to the *sophisticated regularity view*, laws of nature are those true generalizations that either are, or are logically entailed by, the basic statements in the *best* description of the universe; that is, the one that strikes the optimal balance between strength and simplicity. There is no pressure on the best description to entail *all* truths about the universe, because a great many truths are of little to no interest to scientists—for example, my living room walls are green, postal boxes in the United Kingdom are red—and we wouldn’t want to count these as laws anyway. Truths about the universe that are not entailed by the best description will not count as laws of nature.

The idea, then, is that while *all gold spheres are less than 1 mile in diameter* may be true, it won’t be included as, or entailed by, any basic statements of the best description of the universe; that is, the one that strikes the optimal trade-off between

strength and simplicity. The generalization *all uranium spheres are less than 1 mile in diameter*, by contrast, will be entailed by some basic statement, perhaps about quantum mechanics, of the best description of the universe.

Thus, the sophisticated regularity view counts as laws just those regularities that are described, or entailed, by basic statements of a particularly efficient description of the universe. As well as the ability to overcome the gold spheres problem, the sophisticated regularity view enjoys a high degree of continuity with actual scientific practice. Scientists are interested in discovering and articulating the most general and informative facts about the universe and in explaining as much as they can about the universe with as few resources as possible. The sophisticated regularity view takes this observation about scientific practice—that scientists seek strong simple explanations—and makes it constitutive of what it is to be a law of nature.

Whereas continuity with science is generally seen as counting in favor of a philosophical theory of laws, some have expressed concern that the sophisticated regularity view's dependence on balancing *strength* and *simplicity* renders laws *mind-dependent*.

What counts as the best balance between the strength and simplicity of a description would seem to depend on contingent psychological facts about human beings. So, if human psychology were sufficiently different, then the system that struck the optimal strength–simplicity balance would be different, hence the *laws* would be different. However, the laws should not be contingent on human psychology, so the objection goes, so the sophisticated regularity view does not provide an appropriate philosophical account of the laws of nature. The defender of the sophisticated regularity view who does not wish to admit that laws are mind-dependent is thus tasked with developing wholly objective standards for evaluating the strength and simplicity of a description of the universe, but this is no mean feat.

Let us now briefly turn our attention to two further, but closely related, problems for the sophisticated regularity view, which have motivated the development of subsequent views to be discussed.

First, the sophisticated regularity view seems to get the order of explanation between laws of nature and facts about the universe the wrong

way around. According to some philosophers, it is intuitively obvious that scientific laws explain facts about the universe: a dropped vase falls to the floor *because* of the law of universal gravitation; a light ray entering a prism exits the prism at angle θ *because* of Snell's law; and so forth. However, according to the sophisticated regularity view, laws of nature merely *describe* facts such as those concerning falling vases and refracted light rays, and so it looks as if facts about the world explain the laws of science rather than the other way around.

A second, related, problem stems from the fact that had the world been very different from how it actually is, the laws of nature, according to the sophisticated regularity view, would have been different too. Since it is *possible* that the world was very different, laws turn out to be highly *contingent* because the matters of fact that they describe are contingent. But perhaps we think that scientific laws are *necessary*; that no matter how the world was, the laws would be just as they are.

Nomic Necessitation

Both regularity views are what philosophers would call “metaphysically lightweight”—they don't posit the existence of any unusual entities that are unfamiliar to common sense or modern science. This is not to say that they do not clash with common sense or science in other ways; we saw earlier that the simple regularity view bestows law status upon regularities that scientists would not deem laws. And both views arguably fail to capture the necessity and explanatory power of laws, features that some will say are why laws are so important to science in the first place. Furthermore, some philosophers have argued that quantum mechanical phenomena pose problems for the most influential versions of the sophisticated regularity view. The *nomic necessitation* view, to be discussed in this section, is less metaphysically lightweight, but its proponents think that this is justified by its other virtues.

The nomic necessitation view can be understood as agreeing with both regularity views that laws are regularities of the form “All Fs are Gs”; the disagreement lies in what *makes* a regularity a law. As before, the proponent of the nomic necessitation view will want to deem the regularities in

List 1 (plus “all gold spheres are less than 1 mile in diameter”) *nonlaws* and those in List 2 *laws*. To do this, they appeal to the idea that there exist such things as *universals* and *relations between universals*.

What are universals? One could think of universals as those entities in the world to which our predicates refer. So, just as names, such as “Aristotle” and “Napoleon” refer to the individuals Aristotle and Napoleon, *predicates* such as “... is red,” “... is human,” “... is massive,” refer to the universals *redness*, *humanity*, *mass*, and so on. It is often said that individuals *instantiate* universals, where instantiation is a special relation that holds between individuals and universals. Take the example of a red postal box: the individual, the postal box, instantiates the universal, *redness*. Indeed, many different individuals can wholly instantiate the same universal; unlike individuals, universals are multiply located and thereby account for respects of similarity between distinct individuals. Aristotle and Napoleon both instantiate *humanity*; postal boxes and cherries instantiate *redness*; electrons, bowling balls, and black holes instantiate *mass*. In each of these cases, instantiating the same universal accounts for similarity between distinct individuals.

The nomic necessitation view also requires *relations* between universals. More precisely, the view requires *necessitation relations* between universals. To get a sense of what we mean by “necessitation relations between universals,” consider the universals *redness* and *color*. One might think that a necessitation relation holds between these universals because, *necessarily*, anything that instantiates *redness* instantiates *color* too.

According to the nomic necessitation view, a regularity of the form “all Fs are Gs” is a law, if, and only if, there exist the universals *F-ness* and *G-ness* and a special kind of necessitation relation called a *nomic necessitation* relation, which we denote “N,” between *F-ness* and *G-ness*, which ensures that anything that is F is G too.

Nomic necessitation relations, unlike the necessitation relation between *redness* and *color*, hold only contingently. So, if it is a law that all Fs are Gs because N holds between *F-ness* and *G-ness*, it is nonetheless possible that N failed to hold and hence that it was not a law that all Fs are Gs. According to the nomic necessitation view, the

necessity with which laws are associated is weaker than the necessary connection between *redness* and *color*. This is thought by its proponents to be a benefit of the present view because, they think, although the laws of nature are necessary *in a sense*, they are not *absolutely* necessary because we can *conceive of* a world in which the actual laws do not hold (by contrast, we cannot conceive of a world in which something is red and not colored) and conceivability is thought (by some) to entail possibility.

The nomic necessitation view deems “all uranium spheres are less than 1 mile in diameter” a law because, it says, there exists an N-relation that holds between the universals *being a uranium sphere* and *being less than 1 mile in diameter*; “all gold spheres are less than 1 mile in diameter” is not a law because there is no N-relation that holds between the universals *being a gold sphere* and *being less than 1 mile in diameter*.

Since the obtaining or not of an N-relation between two universals is a completely mind-independent matter, the nomic necessitation view, unlike the sophisticated regularity view, is not subject to the objection that the laws are mind-dependent. Furthermore, laws, on this view are said to explain matters of fact, rather than the other way around; it is *because* it is a law that all Fs are Gs. In other words, it is because an N-relation holds between *F-ness* and *G-ness*, that any particular F is also a G; N-relations play a *constraining*, or *governing* role, which explains certain matters of fact.

Finally, laws, on this view, are said to enjoy the appropriate *necessity*; if N holds between *F-ness* and *G-ness*, then it is *necessary* that anything that is F is G too.

The nomic necessitation view is not without its problems. Perhaps the most straightforward source of concern is the fact that the view is not metaphysically lightweight. All of the work is done by *necessitation relations between universals*. But we are well within our rights to be skeptical about the existence of universals, let alone necessitation relations between universals. What, exactly, *are* these things and *are* we really justified in believing in them? Universals and necessitation relations don’t seem like other entities familiar to science. It’s not as if we can discover universals and necessitation relations with a sufficiently

powerful particle accelerator. Rather, they are entities posited by philosophers from the comfort of the armchair. But why think that the philosopher's armchair methods can license belief in such peculiar entities?

Relatedly, we may wonder how necessitation relations are supposed to do the work for which they are invoked by the philosopher. A necessitation relation, *N*'s, holding between the universals *F-ness* and *G-ness* is supposed to make it necessary that all *F*s are *G*s. But why is this so? Why couldn't *N* hold between *F-ness* and *G-ness* and yet there still be an *F* that is not a *G*? As David Lewis, a staunch advocate of the sophisticated regularity view, memorably put the point in opposition to David Armstrong, a pioneer of the nomic necessitation view:

I say that *N* deserves the name of 'necessitation' only if, somehow, it really can enter into the requisite necessary connections. It can't enter into them just by bearing a name, any more than one can have mighty biceps just by being called 'Armstrong.' (1983, p. 366)

Advocates of the sophisticated regularity view and the nomic necessitation view do, however, agree, in one important respect, about the nature of physical properties. Physical properties such as charge, mass, and spin are important when it comes to scientific law because laws of nature seem to articulate regularities that are about these properties. Coulomb's law, for example, says that two distinct instances of charge will exert a force on each other proportional to the magnitude of each charge and inversely proportional to the square of the distance by which they are separated. The sophisticated regularity view maintains that laws, such as Coulomb's law, describe how properties are distributed throughout the universe, and the nomic necessitation view maintains that laws are necessitation relations between universals that determine the distribution of properties throughout the universe—this, in essence, is the disagreement between these views. Where the two agree, however, is in the idea that properties could, in the very broadest sense of *could*, be distributed in any way imaginable. They both agree that there is nothing about the property charge, for example, that determines how instances of charge are

distributed and so nothing about charge that determines that instances of charge behave in accordance with Coulomb's law (similar comments apply to other property-law pairs).

This view of physical properties has been held up by some philosophers as a further source of concern for both the sophisticated regularity view *and* the nomic necessitation view. If physical properties are so *thin* as to impose no constraints on the laws that they enter into, it is hard to get a grip on what physical properties really *are*.

The contingency of the connection between properties and laws opens up the possibility that the property that we call *mass* is featured in Coulomb's law and the property that we call *charge* is featured in Newton's law of universal gravitation. In other words, it is possible that mass and charge swapped theoretical roles, which strikes many as very odd indeed. But worse still, since properties and the lawful behaviors with which they are associated are thoroughly contingent, there is no way of *knowing* which property or properties are playing which roles in our scientific theorizing about the laws—properties and their true natures are rendered irrevocably beyond our ken.

Dispositional Essentialism

Dispositional essentialism takes as its starting point dissatisfaction with the thin view of physical properties shared by the sophisticated regularity view and the nomic necessitation view. Rather than conceiving of natural properties as thin and inert, dispositional essentialism maintains that properties have rich *essences*, which determine the laws that they enter into. So, in the same way that you or I may be said to be *essentially* human (but perhaps only accidentally philosophers) properties are *essentially* certain ways too. *Laws*, then, are not imposed on properties, either as external descriptions or necessitation relations; rather, they emanate from the essences of properties themselves because they concern property essences.

According to dispositional essentialism, a regularity such as "all *F*s are *G*s" is a law if, and only if, it holds in virtue of the *essence* of some property or properties. If it is part of the very nature of the properties *F* and *G* that *all F*s are *G*s, then it is

a law that all Fs are Gs. If, on the other hand, all Fs are in fact Gs but this has nothing to do with the essence of any property, then the regularity is merely accidental and not a law.

Returning to our example, dispositional essentialism maintains that it is part of the *essence of charge* that charged individuals interact in accordance with Coulomb's law; in other words, that distinct charged individuals exert a force on each other proportional to the magnitude of the charges and inversely proportional to the square of the distance of separation. Since this is part of the essence of charge, it is not possible that charged individuals failed to interact in accordance with Coulomb's law. So, the odd result (yielded by the thin view of properties shared by the sophisticated regularity view and the nomic necessitation view) that charge and mass, for example, could swap roles and hence that we cannot ever know the true nature of physical properties, is avoided.

Dispositional essentialism yields a simple response to the gold spheres problem, too. It is of *the essence* of some property or properties that there are no uranium spheres greater than a mile in diameter, so it is a law that there are no uranium spheres greater than a mile in diameter. By contrast, there is no property or properties the essence of which makes it the case that there are no *gold* spheres greater than a mile in diameter, so it is merely an accident that there are no gold spheres greater than one mile in diameter.

There is no mind-dependence problem for dispositional essentialism; it is a completely mind-independent matter whether or not it is of the essence of some property that a given regularity holds, so scientific laws are completely mind-independent phenomena.

The question about *order of explanation*, "Do laws explain matters of fact or do matters of fact explain the laws?" is arguably a vexed one for dispositional essentialism. If one interprets laws as describing the essences of properties, it looks as if matters of fact concerning properties and their essences explain the laws. If, however, one interprets laws as *constituents* of property essences—for example, Coulomb's law is part of the essence of *charge*—then one could maintain that laws, so understood, constrain how properties are distributed throughout the universe and, thereby, explain matters of fact. This way of viewing things is not

so far from the nomic necessitation view—on the latter view, laws are necessitation relations between universals that constrain property distributions; on the former, laws are constituents of property essences that constrain property distributions. On both ways of looking at things, then, laws are metaphysical *constraining* entities that determine property distributions and hence determine and explain matters of fact. The obvious difficulty, then, for dispositional essentialism, and shared with the nomic necessitation view, is that of responding to the demand to say more about how, exactly, the laws are supposed to do this constraining work.

Assuming that dispositional essentialism's laws are up to their task, these laws turn out to be necessary in quite a strong sense. Laws concern the essences of properties, so, it is not possible in any sense for properties and the laws with which they are associated to come apart (*compare*: It is part of your essence that you are human, so it is not possible for you to exist and not be human). Instances of charge must behave in accordance with Coulomb's law, for example, because Coulomb's law holds in virtue of the essence of *charge*. The nomic necessitation view, by contrast, allows that instances of charge may not behave in accordance with Coulomb's law if the relevant necessitation relation is not present.

Philosophers disagree on whether the necessity of the laws yielded by dispositional essentialism is an advantage or a disadvantage. Those who think that it is a disadvantage typically believe that conceivability, or what we can *imagine*, is our best guide to what's possible. They argue that since we can imagine instances of *charge* behaving in a way that violates Coulomb's law, it is not necessary that *charge* is governed by Coulomb's law. Others think that conceivability has nothing to do with what's possible and that the laws' being strongly necessary best explains why laws of nature are so important and useful, features that are hard to make sense of on the assumption that the laws are contingent. Some philosophers have even argued that the necessity of the laws yielded by dispositional essentialism does not go far enough because it still allows for the possibility of *alien* properties and laws—merely possible properties that are not actually instantiated anywhere or at any time and which are governed by

completely unfamiliar laws—and this, they argue, severely compromises the laws' necessity and any potential benefit therein.

New Directions

The philosophy of laws is a burgeoning field, and within the spatial constraints of an encyclopedia entry, it's possible to trace only a limited number of lines through this fascinating area of research. Let us now examine some new research directions before taking a final look at the broad philosophical issues raised in this entry.

Structuralism

The philosophy of scientific laws is entwined with the philosophy of *modality*, that is the philosophy of possibility and necessity. This was seen in the preceding discussion of whether and to what extent the laws are necessary. In general, we may think that laws play a role in determining the limits of what is possible. *Ontic structural realists*, motivated by developments in current physics, argue that the world fundamentally *is* a modal structure. That is to say, there are no individual objects at the fundamental level of the world, just modal structure. And when pressed on what they mean by this, they typically cite laws of nature. So, according to ontic structural realism, such laws are the modal building blocks of our world.

Combination Views

Some philosophers have defended a combination of the sophisticated regularity view and dispositional essentialism. These views combine a broadly dispositional essentialist account of properties, according to which properties have rich essential natures that constrain how they are distributed, with the idea that scientific laws are features of an efficient *description* of property distributions throughout space and time. This combination, they argue, is able to avoid certain problems for regularity views and dispositional essentialism while also reaping benefits from both sides such as continuity with science (from the regularity view component) and the necessity and explanatory power of laws (from the dispositional essentialist component).

Pragmatism

There is a recent trend among some proponents of the sophisticated regularity view to, in a sense, embrace the partial mind-dependence of laws. These philosophers argue that laws ought to be useful tools for creatures *like us*; that part of what it is to be a law of nature is to be something that enables us to make predictions and facilitates the manipulation of our environment to our own ends. Since what counts as useful for us will, to an extent, depend on our *minds*, the laws are partially mind-dependent. But these philosophers think that this is as it should be if we are to provide a sensible philosophical analysis of laws that is genuinely continuous with actual science.

Conclusion

As this entry has shown, there are various issues of broad philosophical interest around which the question "What is a scientific law?" turns. These include the following. *Coherence with intuitions*: To what extent do we wish to accommodate, for example, the intuition that it is a law that all uranium spheres (but not gold spheres) are less than one mile in diameter? *Continuity with science*: To what extent do we want our account of laws to cohere with the scientific conception of laws? *Mind-(in)dependence*: Can we tolerate a degree of mind-dependence of laws if this achieves continuity with science and perhaps other benefits, or must laws be completely mind-independent? *Metaphysical posits*: Are we willing to accept the existence of universals, necessitation relations, property essences, and governing forces if these yield certain benefits such as explanatory power and the necessity of laws? *Modality*: In what sense are the laws necessary and where does this necessity come from?

Samuel Kimpton-Nye

See also Abduction; Conceptual Analysis; Explanation; Induction; Laws of Nature; Logic, Formal and Informal; Metaphysics; Philosophy of Science

Further Readings

Armstrong, D. M. (1999). *What is a law of nature?* Cambridge University Press.

- Berenstain, N., & Ladyman, J. (2012). Ontic structural realism and modality. In E. Landry & D. Rickles. *Structural realism: Structure, object, and causality*. Springer.
- Bird, A. (2007). *Nature's metaphysics: Laws and properties*. Oxford University Press.
- Chakravartty, A. (2007). *A metaphysics for scientific realism: Knowing the unobservable*. Cambridge University Press.
- Demarest, H. (2017). Powerful properties, powerless laws. In J. D. Jacobs (Ed.), *Causal powers* (pp. 38–53). Oxford University Press.
- Dorst, C. (2017). Towards a best predictive system account of laws of nature. *British Journal for the Philosophy of Science*, 70(3), 877–900. <https://doi.org/10.1093/bjps/axy016>
- French, S. (2014). *The structure of the world: Metaphysics and representation*. Oxford University Press.
- Friend, T. (2016). Laws are conditionals. *European Journal for Philosophy of Science*, 6(1), 123–44. <https://doi.org/10.1007/s13194-015-0131-z>
- Hall, N. (2015). Humean reductionism about laws of nature. In B. Loewer & J. Schaffer. *A companion to David Lewis* (pp. 262–277). Wiley. <https://doi.org/10.1002/9781118398593.ch17>
- Hicks, M. T. (2017, May). Dynamic Humeanism. *British Journal for the Philosophy of Science*. <https://www.journals.uchicago.edu/doi/10.1093/bjps/axx006>
- Kimpton-Nye, S. (2017). Humean Laws in an UnHumean World. *Journal of the American Philosophical Association*, 3(2), 129–147. <https://doi.org/10.1017/apa.2017.19>
- Lange, M. (2009). *Laws and lawmakers: Science, metaphysics, and the laws of nature*. Oxford University Press.
- Lewis, D. (1983). New work for a theory of universals. *Australasian Journal of Philosophy*, 61 (4): 343–377. <https://doi.org/10.1080/00048408312341131>
- Loew, C., & Jaag, S. (2020). Making best systems best for us. *Synthese*, 197(6), 2525–2550. <https://doi.org/10.1007/s11229-018-1829-1>
- Loewer, B. (1996). Humean Supervenience. *Philosophical Topics*, 24(1), 101–127.
- Maudlin, T. (2007). *The metaphysics within physics*. Oxford University Press.
- Mumford, S. (2004). *Laws in nature*. Routledge.
- Tahko, T. E. (2015). The modal status of laws: In defence of a hybrid view. *Philosophical Quarterly*, 65(260): 509–528. <https://philpapers.org/rec/TAHTMS-2>
- Van Fraassen, Bas C. (1989). *Laws and symmetry*. Oxford University Press.
- Wilson, A. (2013). Schaffer on laws of nature. *Philosophical Studies*, 164(3), 653–667.

LAWS OF NATURE

Historically, laws of nature have played a role in the explanations or descriptions of order in the universe. As such, understanding laws of nature is an important task in philosophy of science and the sciences. Understanding laws has traditionally pursued two incompatible paths. One path is nomological realism, where laws have their own metaphysical existence and underlie the patterns and regularities of nature. The other path is typically called the *regularity tradition*, where laws are taken to be descriptions of the regularities and patterns that we find existing in nature. There are only regular patterns, but no presumed metaphysical basis, underlying these regularities. Alternative paths seek to understand laws in terms of causes or dispositions, counterfactuals, and constraints. Philosophers usually think of scientific laws—the laws scientists state in their models and theories—as approximations to the laws of nature.

Nomological Realism

Nomological realist accounts posit a metaphysical existence for the laws of nature. Laws are fundamental to reality and are often thought of as *governing* the behavior of the events and patterns we experience in the world because there is a form of necessity at work in the law itself. The laws of nature are conceived as having a kind of metaphysical autonomy, and everything that happens in the universe follows because of the laws. This is a prescriptivist view of laws of nature.

Universals Accounts

Universals accounts posit that laws are grounded in necessary relations among universals. A universal is a class of mind-independent entities such as red or round. Universals are qualities or properties that individuals can share. For instance, a particular ball can be red and round, as can a particular ornament. Universals accounts formulate law statements as “All Fs are Gs,” where the universals F and G cannot be further reduced to anything more basic.

An example under this account would be “All electrons have charge $-e$,” where e represents the

exact magnitude of the charge. Electrons are basic particles that cannot be reduced to any constituent parts and share the universal electron-ness; the magnitude of the charge e cannot be further divided. The law statement is taken to be true because both electron-ness and e are universals where the relation between them is one of necessitation. The law is thought to explain why we always and only see electrons with the same magnitude of charge.

The nature of the necessitation has two possible elaborations. One branch will say that the necessitating relationship between electrons and charge is one of necessary governance: Electrons have the charge $-e$ by virtue of the law of nature. The form of necessity may be further specified as being physical or metaphysical necessitation. The other branch will say that the charge $-e$ necessarily inheres in the electron as a physical or metaphysical necessity. On either branch, laws of nature are prescriptivist.

Proponents argue that nomological realist accounts make sense of counterfactual reasoning that is a staple in scientific investigation. For instance, if we know that “All electrons have charge $-e$ ” is a necessity, scientists can reason through questions such as “If a magnetic field of X configuration were applied to electrons, how would their motion be affected?”

Laws as Primitive

This kind of nomological realist account assumes that laws of nature are primitive or fundamental necessities. Every theory, metaphysical system, or ontology must have something that is unanalyzable and basic. An example would be Euclidean geometry with its five fundamental axioms as unanalyzable primitives. The laws, then, are the basis for explanation for all patterns such as “All electrons have charge $-e$.” As such, it is left unanalyzed whether electron-ness and charge are universals and whether the necessary relationship between them is one of governance or inherence.

Regularity Accounts

Proponents of nomological realist accounts believe there must be some underlying *metaphysical glue* holding the world together. In contrast, regularity

accounts assume there is no physical or metaphysical necessity at rock bottom making the patterns and regularities of the world what they are. Rather, the patterns simply are what they are. Laws of nature, then, are descriptivist—they describe the patterns we find in nature.

It is rare to find anyone defending a view of regularities as simple as this. For one thing, it is not able to distinguish between accidental truths such as “All gold spheres are less than a mile in diameter” from the kinds of regularities that are crucial to scientific explanation like “All electrons have charge $-e$.” Another difficulty is that this simple regularity view runs into problems grounding inductive inferences. Instead, defenders of regularity accounts have developed more sophisticated approaches.

As a historical note, it is common in the philosophy literature on laws of nature to speak of regularity accounts as “Humean,” referring to the Scottish philosopher David Hume (1711–1776). Although Hume believed we come to associate patterns such as cause and effect through our experience of observing events occurring together (e.g., that rocks fall when dropped), he was a firm believer in laws as representing necessities. What Hume questioned was our empirical access to, our ability to know, these necessities. Hence, it is not historically accurate to refer to regularity accounts as Humean accounts.

Best-Systems Accounts

To handle issues with scientific inferences, best-system accounts seek to organize the truths representing regularities into an ordered system that supports deductive inferences modeled on Euclidean geometry. The idea is that laws of nature are those descriptions of generalizations that would appear at the base as axioms in the system. The axioms are chosen to fit the best combination of content, strength, and simplicity. If well designed, the system supports derivation of observations from the laws plus a specification of the context. This is the spirit of how Isaac Newton’s laws of motion applied to particular contexts allow for derivation of motion such as the behavior of a swinging pendulum.

When scientists discover that a regularity they thought was a law actually is not (perhaps because they were mistaken about the regularity or

discover it is a consequence of a more basic regularity), it is possible to adjust the system by replacing the regularity in question with a more appropriate one. Therefore, the existence of a best system does not preclude revising our knowledge on the basis of empirical findings or discovering new laws.

Best-system accounts make sense of scientists' counterfactual reasoning in terms of the structure of the best system by using the knowledge of the patterns functioning as axioms: Scientists derive how a change of magnetic field configuration *X* will affect the behavior of electrons using the laws describing persistent patterns.

Causal/Dispositional Accounts

Some accounts of laws of nature attempt to avoid the metaphysical glue of nomological realism as well as the mere descriptivist nature of regularity accounts. One of these alternatives is to ground the persistent patterns of nature in causal capacities or dispositions. It is these capacities or dispositions that are the truth-makers for laws of nature.

An example would be that electric charge necessarily has the disposition to repel like charges. This disposition underlies the law that particles with like charges repel each other (e.g., two electrons), while particles with opposite charges attract (e.g., electron and proton). Hence, causal or dispositional accounts end up locating some form of necessity in the capacities or dispositions in nature rather than in the laws. There is still a form of metaphysical glue that explains the laws of nature.

Counterfactual reasoning on these accounts would look at how a change in a causal capacity would lead to an observable change. For instance, if a change in a magnetic field is made to configuration *X*, this will change its disposition to affect electron behavior.

Counterfactual Accounts

The need to support counterfactual reasoning about laws of nature and concrete situations appears in all accounts of laws. Hence, another alternative to either nomological realist or regularity accounts is to move directly to laws as

counterfactuals or as underwritten by counterfactuals. Roughly, the idea is that instead of thinking laws of nature support counterfactuals, it is counterfactuals that support laws of nature.

To make sense of counterfactual accounts, we need the concept of possible worlds. Possible worlds are defined as a set of propositions that are true of the entire history of a universe or as the entire set of facts that compose the history of a universe. For example, the total set of facts comprising the history of our universe is the actual world, so it is a possible world. So, too, is a universe where all the facts are the same as the actual one except that you are half an inch taller. So, too, is a universe where all the facts are the same up to Japan's surrender in World War II before any atomic bombs were dropped. In a sense that can be made more precise, the possible world where you are a half inch taller is closer to the actual world than the possible world where Japan surrenders early.

Counterfactual accounts of laws focus on sets of possible worlds that share the same laws of nature as the actual world and then explore variations. Counterfactuals themselves are spelled out in terms of the differences in the possible worlds (e.g., in terms of differences in the propositions comprising the entire history of the different possible universes). What it means to be a law of nature is spelled out in terms of the kind of necessity that can be defined for a set of possible worlds. This necessity, in turn, is grounded in the counterfactual propositions true in these possible worlds. The relevant kind of necessity is not metaphysical or physical necessity, but a modal necessity depending on grades of possibility.

Constraint Accounts

The foregoing analyses focus almost exclusively on the status of law statements such as "All electrons have charge $-e$." However, in actual scientific practice, laws typically refer to equations (usually in the form of differential equations) expressing relations among variables and forces describing the evolution of a system state. Such equations and states show up in models of physical systems. An example would be Newton's second law, $F = ma$, where F represents a mathematical model for the force acting on a particle, a is the

acceleration or rate of change of the particle's velocity, and m is the particle's mass.

Constraint accounts capture these aspects of scientific practice while steering a course between nomological realist or descriptivist understandings of laws of nature is. The modeling of forces and system states in physics generally is carried out in terms of energy and constraints. Such models are defined over a state space, an abstract mathematical space of points where each point represents a possible state of the system. This defines the space of possibilities for the system. An instantaneous state of the system is characterized by the instantaneous values of the variables considered crucial for a complete description.

Laws of nature, on this account, constrain the action of forces along with other constraints of system behaviors. The constraints can be physical, dimensional (e.g., some forces behave very differently in two dimensions vs. three), dynamical (e.g., due to the dynamics of system behavior that arises over time), and so forth. Physicists pay attention to the constraints defining the particulars of a physical context to understand system behaviors. In fluid convection, for instance, the emergent large-scale fluid rotation constrains possible motions of smaller scale fluid molecules under the action of gravity (which itself is constrained by spatial dimension).

The most fundamental laws are those that define the total space of physical possibilities for the universe as a system and its subsystems. Yet, these laws do not dictate the behavior of any systems but constrain those behaviors to be physically possible. Further constraints restrict both the kinds of systems that can exist and their behaviors. Hence, chemical or biological laws and properties constrain system behaviors further to specific regions of possibility space, yet never violate the fundamental laws—never lead to behavior that violates the fundamental constraints of physical possibility.

On constraint accounts, laws of nature are constraints, so are not mere descriptions. On the other hand, no appeals to metaphysics are needed to explain why systems such as the solar system behave as they do. Constraints, forces, and energy in their contingent configurations are primarily responsible for the behavior of systems studied by physicists constrained by fundamental laws plus any laws that emerge over time. Additional

constraints and properties are needed for explanations of geological or biological systems.

Statistical Laws

Statistical laws are very common in the sciences. For example, radioactive decay is a statistical phenomenon. Consider a sample of Uranium 232. In 1 hour, we can say on average how many nuclei will undergo a decay event but cannot say specifically which nuclei in the sample will decay. What we observe (and what physicists derive from physical principles) is that these decay events will conform to a statistical regularity, a statistical law. Scientifically, the randomness of such systems always exhibits a form of order. Order that is different from deterministic order, to be sure; nevertheless, randomness observed by scientists always is statistically ordered.

How do different accounts of laws of nature make sense of statistical laws? Nomological realists struggle with how to make sense of an appropriate form of necessity that yields statistical behavior. One way out of this struggle is to assume that all statistical behavior has an underlying deterministic basis so that all statistical behavior is epistemic only—a lack of knowledge about or empirical access to this underlying basis. Primitivists can simply posit brute statistical laws as basic to the universe.

Proponents of causal/dispositional accounts also struggle with statistical laws. Some accounts try to invoke statistical behavior as a reflection of our lack of epistemic access to an underlying deterministic disposition. Others explore statistical or indeterministic forms of causation as possible explainers for statistical laws.

Similarly, proponents of counterfactual accounts choose between counterfactuals that represent ignorance of an underlying deterministic basis, or between exploring statistical counterfactuals.

On regularity accounts, statistical laws are simply descriptions of statistical patterns observed in nature. Such accounts can remain neutral about whether observed statistical behavior is epistemic or represents randomness that is irreducible to a deterministic basis.

For constraint accounts, statistical behavior depends on the relevant laws constraining system behavior, as well as contingent conditions providing additional constraints that locate concrete

system behavior in either deterministic or indeterministic portions of physical possibility space. The laws as part of the total context play a role in whether behavior is statistical or not.

Robert C. Bishop

See also Explanation; Indeterminism; Induction; Laws, Scientific; Modeling

Further Readings

- Armstrong, D. (1983). *What is a law of nature?* Cambridge University Press.
- Bird, A. (2005). The dispositionalist conception of laws. *Foundations of Science*, 10, 353–370. <https://doi.org/10.1007/s10699-004-5259-9>
- Carroll, J. W. (2020). Laws of nature. In E. N. Zalta (Ed.), *The Stanford encyclopedia of philosophy* (Winter 2020 Edition). <https://plato.stanford.edu/archives/win2020/entries/laws-of-nature/>
- Lange, M. (2009). *Laws and lawmakers*. Oxford University Press.
- Lewis, D. (1986). *Philosophical papers* (Vol. 2). Oxford University Press.
- Koperski, J. (2020). *Divine action, determinism, and the laws of nature*. Routledge.
- van Fraassen, B. (1989). *Laws and symmetry*. Clarendon Press.

LIFE SCIENCES, CONTEMPORARY

Theories and the life sciences enjoy a special relationship. For some, theories about the natural world—explanations of nature at the level of fundamental laws and rigid causality—are still the main goal to achieve, the sign of biology maturing into an exact science in the 21st century. For others, the life sciences are on the brink of a sea change in scientific explanation. Understanding the complexity of organismal life and organization consists not in reducing its complexity to constitutive elements and their causal relation but in the reverse process of simulating the behavior of the entire system and then inferring from the dynamics of the procedure how the system works.

Although space does not permit this entry to offer a comprehensive account of the life sciences or the concept of theory itself, it does aim to

elucidate the place of theory in the contemporary life sciences. In the words of the Italian philosopher, physiologist, and priest Lazzaro Spallanzani, it sheds light on some relevant constellations just as a “mirror . . . changes the appearance of objects” (1803, p. 326).

Nature, Science, and Scientists

The usefulness of theories hinges on the ability to validate them: Theories provide predictions that can be tested. They are analytical tools which subject themselves to examination for logical correctness and consistency; they formulate a set of mechanistic explanations of known facts which lead observed patterns and regularities back to underlying principles; they form a unifying framework which can reconcile and explain a multitude of empirical findings, even contradictory ones; they give biological findings, framed in terms of mathematical formalizations, logical rigor, unambiguousness, and stringency.

In nature, the life sciences have a research topic which is resistant to these forms of examination as practiced in the contemporary scientific theory and philosophy of the exact sciences, above all physics. There is, on the one hand, the subject itself: Nature is diverse and complex. It is the epitome of abundance: a profusion of variety and form, an incalculable diversity of existing organisms, and open to the future. Nature is also both topical and historical as well as coincidental. As such, it bears the past of its evolution within itself, a past which continues the living into a series of possible futures without determining the future or making it predictable. Nature is, after all, *becoming*, not *being*. Living is principally a continuous process, constantly active and permanently in transition. Susceptible to the fleeting details of everyday changes, living beings are part of an environment which they create and which has created them.

That leads, on the other hand, to a host of independent sciences devoted to this nature: botany, zoology, ecology, genetics, evolutionary biology, physiology, cell biology, epidemiology, and paleontology are just a few disciplines united under the umbrella concept of the life sciences. Each constitutes an individual field of know-how, with a distinct body of knowledge, questions, methods, and practices.

Third, and no less important, is the role of the scientist whose expertise is incumbent on the testing of theories and models. In the 21st century, the fragmentation of knowledge and the increasing specialization of scientific research cast a spotlight on the institutional, communicative, and participatory organization of science in addition to the purely technical and substantial competence of individual scientists. The scrutiny of a mathematical model or theory about a biological matter is essentially a question aimed equally at the competence of the biologist, who is expected to understand the mathematical formulation of biological matters, as well as at the mathematician, who must be informed about the biology of what they formulate in mathematical terms. For decades now, ever new courses of study, curricula, and interdisciplinary centers have been established at universities and research facilities around the globe whose aim is to merge the biologist and mathematician into the competences of a single person. In fact, what is being demanded here is a much wider spectrum of specialist knowledge, since the interdisciplinary research structure of the contemporary life sciences often brings all manner of biologists, computer scientists, mathematicians, physicists, and chemists together in one laboratory just as, for example, marine biologists, microbiologists, and ecologists set out together on an Arctic expedition. Declaring the validity of theoretical knowledge is, alongside the need to be commensurate with epistemology and exert due to logical care, thus also a process of the disciplinary negotiation of meaning; achieving consensus on the use of words, formulas, and images in which knowledge is captured; as well as the attribution of authority to the infrastructures, institutions, and spokespersons that produce this knowledge.

History, Philosophy, and Theory

The philosophy and history of science—and particularly *Erkenntnistheorie* or epistemology, the theory of knowledge—have long pondered the relationships and dynamics, analytical as well as historical, of the many ways of thinking and reasoning in the life sciences from their beginnings to the present day.

From a historical perspective, theories become the subject of narratives. Whereas in the past these

were often narratives of progress—in the sense of the increasingly successful theoretical grasping of a subject often associated with hierarchies of value—historical epistemology adopts an approach that treats the sciences themselves as a historical object. As the philosopher Gaston Bachelard noted, the social, cultural, and technical conditions of science in a particular place at a particular time are historically called into question here as the conditions of possibility for the coming and going, transformation and legitimization of research objects, epistemic categories, the science(s) and knowledge *tout court*.

Arguably the most famous theory in the history of biology, Charles Darwin's theory of evolution is an example of a theoretical edifice in various guises. In the 19th century, Darwin's theory of evolution by natural selection presented a mechanism to explain how populations of organisms change over time. Darwin himself called *natural selection* a theory even if he regretted having too little understanding of mathematics to be able to give mathematical form to his idea. This was first achieved by population genetics. Independently of one another, the German doctor Wilhelm Weinberg and the British mathematician Godfrey Harold Hardy formulated the Hardy–Weinberg principle in the early 20th century. Natural selection, according to this basic premise, only occurs where sufficient genetic variation exists in a population. Given an ideal population, the Hardy–Weinberg principle states that genetic variations remain constant under certain circumstances, namely in the absence of selection, mutation, migration, or genetic drift. Although in reality there are no ideal populations, the algebraic formula is nevertheless suitable for making predictions about the relative frequency with which a gene (dominant or recessive) is to be found in an actual population. The British statistician and geneticist Sir Ronald Fisher employed mathematical and statistical methods in his 1930 publication *The Genetical Theory of Natural Selection* and thus provided numerous concepts and models, such as Fisher's fundamental theorem of natural selection, Fisher's principle, and Fisherian runaway selection.

Whereas population genetics concentrates on the genetic occurrences in a population, the mathematical game theory models their evolution in the framework of competing strategies within a

population. The cooperative and competitive behavior of individual protagonists is modeled in a game-theory framework of contests, strategies, and frequencies with which they are applied. In this context, games are the name for mathematical objects which function according to different rules since they formalize varying forms of behavior, strategies, and problems within a population. Another way of formulating natural selection in mathematical terms is via the incidence of mutations; that is, via the dynamics of the population's evolutionary change, as evolutionary invasion analysis, for example, has been doing with the tools of dynamical systems theory since the end of the 20th century.

By contrast, philosophy and the theory of science analyze the various ways of targeting theoretical knowledge about nature in terms of logical rigor and robustness, and the degree of formalization they achieve. Alternatively, they fundamentally question the concept of theory for the biological world. Epistemologically, the arguments revolve around the unique characteristics of life. What is called into question is whether theories, if it is possible to formulate them, can capture the living world with physical, mathematical, and chemical descriptions or just the elements they share with them, but not those which go beyond.

Within the epistemic diversity of theoretical concepts in the life sciences, hypotheses and models occupy a special place. As a rule, the hypothesis precedes the theory: the one taken as the mere suggestion of an explanation, the other as a technical and robust working out of a hypothesis into a theory. Experimentation, on the other hand, the controlled empirical setup for the acquisition of data, is alternatively considered as guided by hypotheses or informed guesswork—rarely taking place in the framework of a fully fledged theory but just as a mere tool for data collection—or itself considered a model system of its own Rheinberger (1997). The term *model* is the most dazzling of all and, according to some observers, currently over-used as Morgan & Morrison (1999). The notion of a model encompasses concepts, materials, practices, computation, or images and also extends to robotics and artificial intelligence. It figures as an intermediary, a bridge between theories and data. Rather than devaluing the term, the

abundance of modeling strategies is valued as the characteristic combinatory style of reasoning in the field following on from the complexity of the biological world.

Data, Dynamics, and Computers

Recent developments in the life sciences add to the epistemic conundrum of the study of nature. Computation, engineering, and data acquisition have been key to advances in the 21st-century life sciences. Most notable is the rise of systems biology, which is the attempt to study life by regarding it in its complexity, examining it computationally, and providing types of explanation different from those previously known in the history of biology. Other developing fields such as synthetic biology continue to build organismal life from scratch. To this effect, they employ a growing and increasingly sophisticated number of technologies and foster epistemologies that equate the artificial *making* of biological worlds with the classic scientific forms of understanding and explanation, if not claiming it a superior one.

Faced with an unprecedented deluge of data, some have been prompted to imagine the end of theory altogether and others to embrace computation and engineering as approaches to rescue biology from observation and experiment, for centuries the pillars of knowledge acquisition about the natural world. Instead, they claim that computation and engineering will finally allow biology to enter a new era of hypothesis and theory-driven research. The computer, especially, is set to open a pathway to new kinds of theoretical knowledge via computational modeling and simulation.

With the large-scale investigation of the dynamics of biological systems and attempts to predict their behavior, the life sciences are finally breaking new ground. Along with this change in perspective comes a revolution in the visual imagery of the life sciences, where computer animations have become ubiquitous and central to working with, making sense of, and communicating, the otherwise impenetrable data.

In order to depict the changing landscape of the life sciences in the post-genomics era, a series of alternative epistemological frameworks has been put forward: Contemporary research is described, by Creager et al. (2007) for example, as “science

without law” as “data-centric” by Leonelli (2016), and by Fox-Keller (2003), as proceeding via “modelling from above” Wise (2004) by contrast calls it “growing explanations.” For scientists in the field, their work is frequently the more mundane practice of “tinkering” and “reasoning for results” as Bray (2001).

Historically rich in diverse practices, concepts, and competing approaches, the life sciences continue to develop manifold ways of thinking and reasoning about the living world that do not necessarily amount to theories in the aforementioned sense. Instead, there are many ways of pursuing theoretical knowledge in the life sciences, which may or may not contribute to a more unified understanding of the biological world at large. Theory in biology, in other words, is conducted in models, visualizations, simulations, and concepts; it is pursued with wet lab and *in silico* (computer-model) experiments; and it is not limited to mathematical formalization but also pays attention to the epistemological foundations of defining knowledge about the baffling wealth of the biological world.

Among the most notable novel tools in this shifting 21st-century scientific landscape are two which are performed on the computer but which have an impact far beyond the realms of science: trial theories or simulations, and computer animations; in other words, moving data.

Computer Simulations

Computer simulations emerged in the context of physics and the Second World War but have grown to become a ubiquitous tool in practically all fields of science over the last half-century. Having only come to the fore in the life sciences at the turn of the new millennium, they make it possible to study aspects of nature which were less accessible or completely inaccessible with the previous repertoire of techniques, tools, and practices. First, simulations target the complexity of the biological world; second, they investigate the living world as a dynamic system; third, they aim to predict the future behavior of large systems with the goal of intervening in and, if possible, controlling the developments to come.

Simulations are a tool of the moment, which reflects the interdisciplinary reality and dynamic

alliances of scientific and technological expertise in life science research. Their power is that “glue” and “compromise” Sismondo (1999) due to their ambivalent nature they create new cross connections to long-established scientific demarcations: They not only cut across disciplines but also across analysis and prediction, understanding and doing, and science and technology. They thus permit the integration of previously heterogeneous bodies of knowledge.

That hybridity of simulations compels science to traverse its narrow academic boundaries and become an actor in the world in hitherto unseen ways. Because they address real-world problems on a large scale, such as the evolution of entire ecosystems or climate change, and are designed to predict their development, thereby revealing ways to intervene in the course of processes, simulations are not only in themselves a scientific object in motion (technically working through constant algorithmic iteration), they also create shifting webs of activities spanning from scientists to activists.

In stark contrast to the reductionist tradition of taking things apart, which has characterized biology for centuries, a field such as systems biology integrates data from the microlevel of cells and proteins into organs and entire organisms to tackle biological complexity.

A whole range of simulation techniques are deployed in the life sciences. *Monte Carlo simulations*, based on random numbers and stochastics, can be used, for example, to model intracellular molecular transport; *agent-based modeling* focuses on the rule-based interaction of independent agents (individual or collective) and can be applied to investigate the behavior of—for example—animals in populations interacting with each other and the environment; *cellular automata* have a long history in biology and model the dynamics of systems and their various states dependent on their neighbors and are thus useful in studying biological pattern formation or the motion of fluids at the microscopic level; *neural networks* are part of the field of machine learning in which the instructions, given in the form of algorithms, are inspired by the neural network of the brain; and *finite-element simulations* are an engineering method to calculate the deformation of given structures by discretizing them and solving a set of differential equations. In

embryology, for example, they are employed to understand the development of morphological structures in ontogenesis by calculating shape changes in early gastrula formation.

Operating on the level of entire systems and their behavior, simulations enforce the study of the living world in its temporal unfolding. They frame organisms as beings in time, processing information in time and constantly integrating various levels of time including the temporal regimes of nanoseconds in cells and subcellular structures, diurnal rhythms, or life cycles and evolution.

The dynamics of life come to the fore in simulation techniques because the iterative structure of simulations mimics the processual character of nature. Put simply, computer simulations are mathematical models. In order to run them on a computer, some models are discretized versions of the mathematical laws assumed to rule the natural process in question. Other models implement no physically or chemically significant algorithms but reproduce the behavior of a system on a phenomenological or statistical basis. In the computer, a range of dynamic phenomena become analytically tractable that have previously been out of reach of mathematical formalization and its main tool, differential equations. Once implemented, those models are run on the computer, and one can follow how the iteration of the same set of rules again and again produces some kind of outcome or rather a range of outcomes. It therefore follows, as noted by Stephan Hartmann, that simulations work by way of one process imitating another.

Initially, simulations were employed to test hypotheses, guide the design of an experiment, explore and tune the range of parameters, or fit the experimental setup optimally to the material and intellectual resources available. Increasingly, however, they are used to predict the outcome of a process. Prediction is considered a prime element of theory; simulations, however, are applied in fields of biological inquiry where there is no theory. Simulations have therefore been lauded as trial theories and computation, previously a tool to apply theory, as the novel agent of building theories.

Scientific Animation

Imaging technologies are vital to the acquiring and processing of data, the analysis of simulation

results, and the display of dynamic biological features. Revolutionary advances such as live cell imaging allow us, for example, to trace the processes of meiosis and mitosis, see molecular motors transport organelles in the cytoplasm, or follow leukocytes on their way to the site of inflammation in their temporal dynamics, *in vivo*, in a live specimen.

Even if they are generated by technical means—including fluorescent genes, high-resolution laser microscopes, and magnetic resonance—the imagery of contemporary life sciences is, on the one hand, natural images in the sense of images of nature. As computer-generated and animated technical images, they create a sensory perception of processes in nature, and it is this visualization that opens up a new understanding of nature as we regard its movements and processes. On the other hand, these images are also an experiment, a mathematical model, and an algorithm. With the aid of technical tools, they produce a dynamic manifestation of nature to which the human sensory apparatus has no access. In addition, the numerous imaging technologies serve in the first place as a vehicle to digitize information. Data of subcellular processes, generated with the aid of a microscope, are no longer seen and analyzed by a human viewer but registered and further processed by a computer program. The images are therefore the visual result of numerous conversions between experiment and observation, organism and computation, and the material and the virtual realm. With the imaging of living organisms under life conditions, experimenting with the living body, building models, and imaging organic processes as they proceed in time are performed in unison. Previously, separate analytical categories thus collapse into a single domain of knowledge, with unknown consequences for the future conduct of research.

Extremely complex processes, such as embryogenesis, the development of a new organism from a single cell, are a field in which the approach of systems biology—studying the behavior of a system as a dynamic whole—constitutes both a major challenge and a radical turning point in research culture. The process of formation is a coordinated progression of spatiotemporal, synchronized cell movements which can be observed in the computer simulation as a direct motion sequence. *In vivo* imaging technologies like those

described above track every single cell from every angle, moment by moment. Systems approaches thus make it possible to generate massive sets of data. These are the prerequisite for the modeling of development in computers. If the data are sufficiently detailed, the research will lead to a computer model of embryogenesis that is able to make predictions about the behavior of the system and consequently intervene in the various parameters of this process at any time and place whatsoever.

The animated images of embryonic development are not just digital visualizations of a process. The embryos are no longer simply depicted in the trajectories of fluorescent cells on a dark screen. What we see is not merely development in four dimensions, a highly complex process of coordinating and differentiating a single cell into its thousands of successors. The embryos are now themselves a computer program, and their development can be studied and controlled like the course of software on a computer. In this respect, the *in vivo* visualizations bear no resemblance to the traditional depictions of development in embryology.

As a science of syntheses and relationships, embryology has been concerned with complexity, integration, and context from the very outset. With the aid of the computer, it is now no longer the aim of research to fix embryos histologically, manipulate them experimentally, and then make their development visible in a series of images, as was the practice for centuries. Instead, research now observes how organ formation advances in numerous steps from a single cell to more complex structures. To this end, a model of development based on experimental data is run on a computer, the complexity of formation increasing with every iteration, as does the amount of experimental data. Becoming more complete and data-saturated with every run and modification, the model is supposed to become increasingly predictive, too. The *in silico* model is thus aimed at serving as a hypothesis of how the process works as a whole, *in vivo*. Working with individual components and their relationships, development is framed as an emergent property whose qualities can be understood not through analysis of the components but through the integration of data from the level of cells to the entire organism and its maternal environment. Edelman et al. (2010) and Khairy and Keller (2011) explain all of this thoroughly.

Conclusion

Making the dynamics of nature visible and the moving image a topic of research changes our view not only of nature as a whole but also of the individual. The modeling of processes is based on large sets of data and follows the rules of mathematical and computational formulization. Experiments, on the other hand, are still conducted on individual organisms and with the aid of individual model organisms. The dynamization of nature in systems biology and scientific animation is thus based on both individual surveys and statistical values. This raises the fundamental question of whether the processes and events that constitute nature have a meaning as singular occurrences or whether future science will redefine the individual as merely an average value within a possible sample space.

Janina Wellmann

See also Biology, Evolutionary Epistemology; Hypothesis Testing; Modeling; Simulation

Further Readings

- Bachelard, G. (1986/1934). *The new scientific spirit*. Beacon Press. (Original work published 1934 as *Le nouvel esprit scientifique*. Quadrige).
- Bray, D. (2001). Reasoning for results. *Nature*, 412(6850), 863–863. <http://doi.org/10.1038/35091132>
- Creager, A. N., Lunbeck, H., Wise, E., & Norton, M. (Eds.) (2007). *Science without laws: Model systems, cases, exemplary narratives*. Duke University Press.
- Edelman, L. B., Chandrasekaran, S., & Price, N. D. (2010). Systems biology of embryogenesis. *Reproduction, Fertility, and Development*, 22(1), 98–105. <http://doi.org/10.1071/RD09215>
- Fox-Keller, E. (2003). Models, simulations, and “computer experiments”. In H. Radder (Ed.), *The philosophy of scientific experimentation* (pp. 198–215). University of Pittsburgh Press.
- Galison, P. (1996). Computer simulations and the trading zone. In P. Galison & D. J. Stump. (Eds.), *The disunity of science: Boundaries, contexts, and power* (pp. 118–157). Stanford University Press.
- Hartmann, S. (1996). The world as a process. In R. Hegselmann, U. Muller, & K. G. Troitzsch. (Eds.), *Modelling and simulation in the social sciences from the philosophy of science point of view* (pp. 77–100). Kluwer Academic Publishers.

- Khairy, K., & Keller, P. J. (2011). Reconstructing embryonic development. *Genesis*, 49, 488–513.
- Leonelli, S. (2016). *Data-centric biology: A philosophical study*. University of Chicago Press.
- Morgan, M. S., & Morrison, M. (Eds.). (1999). *Models as mediators. Perspectives on natural and social science*. Cambridge University Press.
- Rheinberger, H.-J. (1997). *Toward a history of epistemic things: Synthesizing proteins in the test tube*. Stanford University Press.
- Shou, W., Bergstrom, C. T., Chakraborty, A. K., & Skinner, F. K. (2015). Theory, models and biology. *eLife*, 4, e07158–e07158. <http://doi.org/10.7554/eLife.07158>
- Sismondo, S. (1999). Models, simulations, and their objects. *Science in Context*, 12(2), 247–260. <http://doi.org/10.1017/S0269889700003409>
- Spallanzani, L., Dalyell, J. G., & Bonnet, C. (1803). *Tracts on the natural history of animals and vegetables* (Vol. 1, 2nd ed.). Printed for W. Creech and A. Constable.
- Strasser, B. J. (2019). *Collecting experiments: Making big data biology*. University of Chicago Press.
- Wise, M. N. (Ed.). (2004). *Growing explanations. Historical perspectives on recent science*. Duke University Press.

LINGUISTIC FRAMEWORKS

Linguistic analysis utilizes linguistic frameworks in a variety of ways. Sometimes referred to as *linguistic methods or linguistic approaches*, linguistic frameworks can be understood as conventions, or sets of rules, that enable discussion and analysis of new entities; in other words, they are forms of language that govern the ways in which these entities are described and referenced.

A linguistic framework is thus a way of organizing human communication about wide sets of experiences or observations. A simple example of such a linguistic system would be a mathematical system, in which a problem is solved by deduction from the axioms of the system.

According to the philosopher Rudolf Carnap (1891–1970), for every linguistic framework, there exist two specific types of questions: internal questions, which are asked and answered within the framework, and external questions asked of

the larger system within which the entities are supposed to exist. Not only did this conception play a central role in logical positivism, but it also greatly influenced the development of the philosophy of science. Both internal and external questions are essential in the building process of *linguistic frameworks*.

Building and Implementation of Linguistic Frameworks

The process of building a linguistic framework can be described as follows. If someone wishes to speak in their language about a new kind of entity, they have to introduce a new system of ways of speaking, subject to new rules. This procedure is called *the construction of a linguistic framework* for the new entities in question. Thereafter, it is necessary to consider whether the entities exist within this framework. This is the internal question. Regarding the external question, it concerns the existence of the entire reality of the system of entities as a whole. Internal questions and answers to them are formulated with the help of the new forms of expressions. The answers may be found by purely logical methods or by empirical methods. This depends on whether the framework is a logical one or a factual one.

The simplest example of entities used in everyday language could be the spatiotemporally ordered system of observable things and events. Once we accept the language with its framework, we can ask empirical questions: *Is there a white paper on the table? Did King Arthur actually live? Are unicorns and centaurs real or imaginary creatures?* These questions need to be answered through empirical investigations. The results are evaluated according to the rules that confirm or refute the evidence for possible answers. For this purpose, explicit rules of evaluation have to be put down in advance.

By contrast, *What are things?* is an external question. The distinction is perhaps even easier to see when the framework in question is that of numbers; clearly *What is 2+2?* is internal and *Are numbers real?* is external. For Carnap, such external questions are not just unanswerable but fundamentally flawed in their formulation. It is

meaningless to ask what is *real* outside the context of the framework because it is the framework that contributes to the cognitive content with which these questions are posed.

Given that natural language emerged as a means of enabling humans to effectively communicate with each other, and not as a means of *finding truth* or any other such modern abstraction, it is reassuring to find a theory that places its emphasis on linguistic convenience rather than on a search for deep metaphysical insights. One might even take a speculative step further to suggest that human minds, primed as they are for linguistic development, are inherently predisposed to accept certain frameworks of exceptional utility (e.g., thing-language) regardless of how flawed or incomplete they might be. It has been suggested that human brains are naturally sensitive to certain grammatical forms and constructions; it is likewise known that even infants have a sense of object persistence and number conservation. Both observations lead one to the idea that thing- and number-language are *natural* frameworks to accept natural ways of organizing our perceptions of the world.

Carnap established the concept of *linguistic frameworks* to cope with the problem of constructing abstract entities within an empiricist frame of reference, and it is not surprising that a major advantage of the idea lies in this area. References to abstractions, particularly those derived from the sciences—coordinates, vectors, and many more—can be freed from implicit Platonism and given their own theoretical grounding as entities introduced by useful and productive linguistic frameworks. This structure offers a clean way of viewing abstractions within an explicitly empiricist worldview.

The linguistic framework thesis is also characterized by a second class of advantages related to traditional ontological problems. One of them is that it does not trip over the nonexistent referent problem; denying the existence of an entity is as simple as stating that the linguistic framework in which the entity is introduced is not productive or expedient.

A related advantage is that the acceptance of linguistic frameworks is fundamentally a pragmatic matter, ultimately adjudicated by *their*

efficiency as instruments. This attitude admits the utility of limited descriptive languages, such as that of Newtonian mechanics, without appealing to the nonsensical concept of *ultimate truth or falsity*.

Instances of Linguistic Frameworks

Linguistic frameworks are used in a number of fields of study, such as grammar (including word classes, categories, morphology, and syntax), graphology, lexis, phonology, pragmatics, and semantics. More specifically, there are seven main language frameworks: *lexis*, the vocabulary of a language; *semantics*, the study of how meaning is created through words and phrases; *grammar*, which includes the system of rules that governs how words and sentences are constructed, the system that groups words into classes according to function, the system of rules about how words function in relation to each other (syntax) and individual units that make up whole words (morphology); *phonology*, the study of sounds in a given language; *pragmatics*, or language in use, meaning how social aspects and intentions affect people's language choices; *graphology*, the study of the appearance of the writing; and *discourse*, the whole piece of text in which more than one sentence is analyzed together with its contexts, info-graphic, audiovisual, and extralinguistic features).

Applications and Usage

Linguistic frameworks have further applications. They are used in the philosophy of language, metaphysics, and philosophy of science. They characterize a chosen language or a set of analytical principles that provide the method and criterion for the formulation of any significant assertion and its solution within the framework. A body of significant knowledge can only be justified by reference to the principles or rules that make up the framework. More precisely, linguistic frameworks enable the operation and logical systematized categorization of the phenomena and concepts within the fields of study and help those fields make references to the real world. The framework could be construed as the basis

for reaching an agreement about any disputed problem.

Different frameworks reflect different ways of talking about the world. But once we have adopted a framework, we must obey all its rules and principles. It could be assumed that any answer to a question about the kinds of entities conceptualized by a language is relative to a framework.

Linguistic Frameworks Used in Linguistic Annotation

Linguistic frameworks are most widely used in linguistic annotations. *Linguistic annotation* signifies any descriptive or analytic notations applied to raw language data. The basic data may be in the form of time functions—audio, video, and/or physiological recordings—or it may be textual. The added notations may include transcriptions of all sorts (from phonetic features to discourse structures), part-of-speech and sense tagging, syntactic analysis, *named entity* identification, coreference annotation, and so on. Although there are several ongoing efforts to provide formats and tools for such annotations and to publish annotated linguistic databases, the scarcity of widely accepted standards has emerged as a critical problem since the late 20th century. Some proposed standards have focused on file formats, while other scholars have argued that standards should instead be focused on the logical structure of linguistic annotations and that researchers thus need to analyze a wide variety of existing annotation formats and demonstrate a common conceptual core: the annotation graph. This would provide a formal framework for constructing, maintaining, and searching linguistic annotations while remaining consistent with many alternative data structures and file formats.

Information Granularity in Applying Linguistic Frameworks

The framework must provide means to represent all varieties of linguistic information (and possibly also other types of information). This includes representing the full range of information from the very general to the information at the finest level of granularity.

In addition, the framework must handle all potential media types, including text, audio, video, and image, and should provide common mechanisms for handling them. The frameworks rely on existing or developing standards for representing multimedia.

Descriptors, information categories, principles, and standards should be fixed in a formal semantic way. The frameworks should also provide common procedures for input interpretation and output generation.

Fields That Apply Linguistic Frameworks

Because linguistic frameworks cover a wide range of systems that pertain to different spheres, a number of related disciplines make extensive use of their resources. The following fields are most commonly noted in this respect.

Computational linguistics is the study of language using the techniques and concepts of computer science, especially with reference to the problems posed by the fields of machine translation, information retrieval, and artificial intelligence. Computational linguistics uses extensively syntactic rules, syntactic trees, and procedures for programming purposes; hence, the usage of linguistic framework is an indispensable part.

Neurolinguistics is the study of the neurological basis of language development and use in human beings, especially of the brain's control over the processes of speech and understanding. Consequently, neurolinguistics constantly refers to semantic, syntactic, lexical, and pragmatic frameworks.

Psycholinguistics is the study of the relationship between linguistic behavior and the psychological processes to underlie it. The processes may involve memory, attention, thought, and others. Hence, grammar, semantic, pragmatic, and discourse frameworks often lend to psycholinguistic studies.

Sociolinguistics studies the interactions between language structure and functioning of society. Pragmatic, discourse, lexical, and phonological frameworks are often used in sociolinguistic investigations.

Statistical linguistics studies the quantitative properties of language. Usage frequencies of acceptable or unacceptable linguistic structures, as well as compiling corpora, are major objects of investigation. Therefore, the entire stock of linguistic frameworks is made use of at one point or another.

Mathematical linguistics studies mathematical properties of language, using concepts from such fields as algebra, computer science, and statistics; hence, it makes wide use of syntactic, phonological, and other linguistic frameworks.

Anthropological linguistics is the study of language variation and use in relation to the cultural patterns and beliefs of humankind, as investigated using the theories and methods of anthropology and referring to lexical, phonological, and discourse frameworks.

Applied linguistics is the application of linguistic theories, methods, and findings to the elucidation of language problems that have arisen in other domains. The term is especially used with reference to the field of foreign language learning and teaching, but it applies equally to several other fields, such as stylistics, lexicography, translation, and language planning, as well as to the clinical and educational fields. Since it covers a wide array of issues, it makes use of the entire linguistic framework.

Biological linguistics is the study of the biological conditions for language development and use in human beings, with reference both to the history of human language and to child development. Syntactic, phonological, morphological, and lexical frameworks often lend to biolinguistic studies.

Clinical linguistics is the application of linguistic theories and methods to the analysis of disorders of spoken, written, or sign language. Any linguistic framework may contribute to clinical studies depending on the topic of research.

Educational linguistics is the application of linguistic theories and methods to the study of the teaching and learning of a language, especially of a first language, in schools and other educational settings. Due to its wide range of issues under investigation, the field of educational linguistics refers to any linguistic framework depending on the circumstances of any particular study.

Ethnolinguistics is the study of language in relation to ethnic types and behavior, especially with reference to the way social interaction proceeds. Discourse, lexical, and pragmatic frameworks are of particular interest, but not exclusive to the field.

Geographical linguistics is the study of the regional distribution of languages and dialects, seen in relation to geographical factors in the environment. Phonological and lexical frameworks are particularly important to such studies.

Philosophical linguistics is the study of the role of language in the elucidation of philosophical concepts, and of the philosophical status of linguistic theories, methods, and observations. Pragmatic and discourse frameworks fall within the realm of these studies.

The overall field of *linguistics* itself provides an invaluable conceptual framework to a number of disciplines that cross-feed from its studies, findings, sources, concepts, and rules.

Mariam Orkodashvili

See also Analysis; Artificial Intelligence; Bioinformatics; Biology, Evolutionary; Cognitive Science; Conceptual Analysis; Cybernetics, 20th Century; Data Models; Evolutionary Psychology; Framework; Hypothesis Testing; Informatics; Information Theory; Linguistics, Contemporary; Modeling; Neuroscience; Paradigm; Philosophy of Science; Pragmatism; Semantics (Introduction to Theory); Semantics (Scientific and Empirical); Statistical Inference, Bayesian; Statistical Inference, Frequentist; Statistics; Syntax (Introduction to Theory); Syntax (Scientific and Empirical)

Further Readings

- Aarts, B. (2007). *Syntactic gradience. The nature of grammatical indeterminacy*. Oxford University Press.
- Aarts, B., Denison, D., Keizer, E., & Popova, G. (Eds.). (2004). *Fuzzy grammar*. Oxford University Press.
- Biber, D., Johansson, S., Leech, G., Conrad, S., & Finegan, E. (1999). *Longman grammar of spoken and written English*. Pearson Education Limited.
- Boeckx, C. (2006). *Linguistic minimalism. Origins, concepts, methods, and aims*. Oxford University Press.
- Carnap, R. (1950). *Empiricism, semantics and ontology*. Massachusetts Institute of Technology Press.
- Carnap, R. (1991). *The philosophy of science*. Massachusetts Institute of Technology Press.
- Chomsky, N. (1957). *Syntactic structures*. Mouton.
- Chomsky, N. (1981). *Lectures on government and binding*. Foris.
- Chomsky, N. (2007). *On language*. The New Press.
- Crystal, D. (Ed.). (1995). *The Cambridge encyclopedia of the English language*. Cambridge University Press, 1995.
- Crystal, D. (Ed.). (1997). *The Cambridge encyclopedia of language*. Cambridge University Press.
- Dancygier, B. (2006). *Conditionals and prediction. Time, knowledge and causation in conditional constructions*. Cambridge University Press.

- Haegeman, L. (2005). *The syntax of negation*. Cambridge University Press.
- Halliday, M. A. K., & Hasan, R. (1976). *Cohesion in English*. Routledge.
- Heusinger, K. von. (2002). Specificity and definiteness in sentence and discourse structure. *Journal of Semantics*, 19, 245–274.
- Johnson, K. (2008). *Quantitative methods in linguistics*. Blackwell.
- Karttunen, L. (1976). Discourse referents. In J. McCawley (Ed.), *Syntax and semantics 7: Notes from the linguistic underground*. Academic Press.
- Quirk, R., Greenbaum, S., Leech, G., & Svartvik, J. (Eds.). (2000). *A comprehensive grammar of the English language*. Pearson Education Limited.
- Widdowson, H. G. (1996). *Linguistics*. Oxford University Press.
- Woods, A., Fletcher, P., & Hughes, A. (1986). *Statistics in language studies. Cambridge textbooks in linguistics*. Cambridge University Press.

LINGUISTICS, CONTEMPORARY

Contemporary linguistics stands at the crossroads of multiple disciplines. It juxtaposes the interests and research questions of diverse fields that might seem unrelated at a glance. However, human language is what provides common research ground for these disciplines. Therefore, contemporary linguistics is the discipline that studies human language. Language of course figures centrally in our lives. Humans discover their identity as individuals and social beings when they acquire language in childhood. Language serves as a means of cognition and communication. It helps us produce thoughts and cooperate with other people. It provides for present needs and future plans. Language seems to be an essential feature of humanity, which enables individuals to rise above the condition of other living being and distinguishes humans from other species. That is why language is the study object of various disciplines.

Language Features

One of the primary and significant features of a language is *arbitrariness*: The forms of linguistic signs bear no natural resemblance to their meaning. The link between them is a matter of convention,

and conventions differ radically across languages. However, there is an exception. Onomatopoeic words resemble to a degree to the sounds that they designate (e.g., Woof! Grrr!).

Another important feature of language is *duality*. Human language operates on two levels of structure. At one level are elements that have no meaning in themselves but which combine to form units at another level that do have meaning. For instance, letters combine in different ways to form words that have meanings. The duality feature operates both on spoken and written levels and indicates the flexible nature of a language.

A further interesting feature of a language is that it enables us to categorize reality and express the categories through language. This important function and feature that language has is *ideational* function. Through ideation function, humans are able to classify and categorize the reality, the environment around them.

Displacement is the next feature of language. It enables speakers to refer to the objects and events that are distant in space or occurred in another time.

Language is meaningful and understandable to other speakers, hence, its *semanticity* feature.

Through a limited number of symbols, language produces innumerable sentences, hence, its *generativity* or *productivity* feature.

Language has rules that govern how symbols can be arranged. These rules allow people to understand messages in that language even if they have never encountered those messages before. Language is not a continuous flow of sounds but a systematically and discretely arranged unity of separate sounds/phonemes. Thus, language has the feature of *discreteness*.

Linguistic Relativism Versus Language Universal

Perhaps the most widely known debate pertaining to *contemporary linguistics* is that between relativist and universalist standpoints. While relativists claim that each language is the reflection of the unique culture and mentality of the speakers of the given language (hence, linguistic relativity), the universalist standpoint (the well-known theory attributed to Noam Chomsky) argues that there is an innate linguistic capacity in all humans to acquire and use the basic rules and structures (parameters)

of any language (hence, the concept of *universal grammar*). Linguistic relativism is also known as the *Whorf hypothesis* or *Sapir-Whorf hypothesis* (named after the linguists Benjamin Lee Whorf and Edward Sapir), which is the view that language shapes cognition; that is, concepts and ways of thinking depend on language. People who speak significantly different languages, then, view the world differently. Moreover, there exists the concept of *linguistic determinism*, which reflects the strong version of relativism and implies the idea that language and its structures limit and determine human knowledge or thought, as well as thought processes such as categorization, memory, and perception. The term implies that people of different languages have different thought processes.

While relativists blame the universal standpoint for restricting the immense capacities of any language to mere structures and rules, universalists defend their stance by further explaining that it is exactly through an innate capacity to acquire the basic rules of any language that humans are able to produce innumerable lexical and grammatical units and to make them comprehensible to speakers of other languages. In fact, in modern linguistic thought, there are many intersecting points between the relativist and universalist positions since the arguments on both sides actually complement, rather than contradict, each other.

Functions of Language

It is noteworthy that from a universalist standpoint, language is *cognitive*. It is seen as a psychological phenomenon that reveals the human mind through the linguistic forms that it produces. However, the cognitive aspect is not the only feature of a language. In order to serve a societal function, it acquires *social* features as well. Hence, it possesses an important communicative function that enables humans to transfer and receive units of information among each other.

Dual Nature of Language and Linguistic Oppositions

As already mentioned, language is an intrinsically dual system. One such duality was identified and put forward by Ferdinand de Saussure, a Swiss linguist, in the beginning of the 20th century. By

duality, he meant the *langue* and *parole* components of a language, by which *langue* signifies the coded system of language rules and principles common to every speaker of that language, and *parole* is the idiosyncratic speech pattern of every individual of that language community.

The dual nature of language is noted by Noam Chomsky as well, who distinguishes between *competence* and *performance*. Closely similar to Saussurean *langue* and *parole*, Chomsky calls the rules and procedures that are commonly acquired by all speakers as *competence*, and the individual speech behavior of any individual as *performance*.

Another significant duality that Saussure put forward was that of the ideas of *synchrony* and *diachrony* in linguistic studies. Synchronic analysis studies the features of a language at the present moment, while diachronic study traces the development and evolution of language throughout centuries by identifying the linguistic changes that the levels of language system undergo. The changes and transformations are traced on all levels of language: phonological, phonetic, morphological, syntactic, lexical, semantic, and even pragmatic (which includes discursive and extralinguistic features).

Dual nature applies to principles of linguistic analysis as well. When conducting linguistic analysis on any level, linguists widely implement the *syntagmatic* and *paradigmatic* dimensions of analysis, *syntagmatic* revealing the horizontal, or linear, relations between the linguistic elements, and *paradigmatic* arranging the elements in a vertical pattern.

The opposition of *markedness* and *unmarkedness* is yet another linguistic feature that characterizes any language. The simplest and most explicit example to illustrate this duality is the existence of negative affixes, lexical units, or syntactic constructions alongside the positive ones (e.g., masked-unmasked, yes-no, went-did not go). While positive linguistic items are considered to be *unmarked*, negative units and constructions are usually considered *marked* during linguistic analysis.

Levels of Linguistic Analysis

Linguistic analysis of language is usually conducted on several levels, which are often closely interconnected and influence each other. Following are the levels that linguistic principles of analysis are

applied to *phonetics* (speech variations), *phonology* (sound variations, phonemes, and allophones), *morphology* (inflections and affixation that bear grammatical value in every language, allomorphs), *syntax* (sentence and phrase structures and means of constructing cohesive and coherent connections within and between sentences), *semantics* (analyzing variations in meanings of lexical or phraseological units), and *pragmatics* (the intentional use of certain linguistic means by speakers to achieve the desired communicative goals). The higher level of analysis goes as far as discourse, which entails longer textual, visual, auditory, graphical, and other extralinguistic elements that influence the overall coding and decoding of the entire speech act.

The study of allophonic manifestations—that is, how the sounds of speech are actually made—is the topic of research of phonetics. The study of phonemes and their relations in sound systems is the research topic of phonology. But these studies need to be seen as interrelated since their findings complement each other. Furthermore, the sounds organize in larger segments called *syllables*, which in their turn are stressed or unstressed. The fact causes the modification of sounds. Besides, sentences are pronounced with rising or falling intonation. All these issues fall under the study of phonetics and phonology.

The next stage of analysis is that of morphology, which studies inflections and word affixations that form grammatical categories and variations of tense, person, case, number, and so on. These derivative forms are called *allomorphs*. Therefore, word derivation, affixation, and inflection are studied on the level of morphology.

Moving on to the next stage of analysis, the syntactic level, the organization and connections of words into larger units are researched. The words organize themselves in phrases, sense groups, and sentences. The sentences organize into even larger units, paragraphs, textual chunks, or larger segments. The main organizing principle is the establishment of cohesive and coherent connections between the words and sense groups through various linking devices, such as conjunctions, adverbs, and so forth. In addition, sentences are studied through the analysis of their immediate constituents, and through the structures of syntactic trees that present building blocks of syntax and computational linguistics.

Morphological and syntactic analyses fall under the grammatical level of analysis, which is primarily concerned with identifying, categorizing, and understanding the structural features of a language. However, these levels of analysis are intrinsically intertwined with other levels of analysis, including the phonetic, phonological, semantic, and discursive levels.

In addition to structural levels of analysis, linguists conduct analysis at the semantic level, which is primarily concerned with the meanings of linguistic units. Componential analysis is implemented for the purpose of understanding the meaning of a linguistic unit. The elements of meaning are analyzed through semantic components, lexemes that carry denotational (and sometimes connotational) sense. Further sense relations between words and phrases are established through synonymy, antonymy, metonymy, polysemy, and so forth.

It is well established that all the aforementioned structural or semantic levels of analysis are indispensably connected with the context and extralinguistic parameters in which these structural or semantic elements are implemented. Hence, pragmatic analysis, discourse analysis, and intertextual analysis are often conducted in conjunction with each other.

The most widely used form of pragmatic analysis is the speech act method, which analyzes linguistic features in the process of speech or communication in which individuals are involved. According to pragmatic analysis, any statement or utterance produced during a speech act engenders three effects: locution, illocution, and perlocution. *Locution* is the expression of the statement itself, *illocution* is the intention that the speaker wants to achieve through making this statement, and *perlocution* is the effect that is made on another person toward whom the illocutionary force was directed. Perlocution could reveal how much the listener got persuaded, impressed, encouraged, or frightened by the speaker: What was the effect of the speech act? How successfully was it conducted? (Hence, the concept of felicity, or success, condition that is often referred to in this respect).

Further, it is often noted by linguists (especially pragmatists) that context is a *schematic* construct. It means that when people make an indexical connection, they do so by linking features of the

language with familiar features of their world, with what is established in their minds as a normal pattern of reality, or *schema*. It is not *out there*, so to speak, but in the mind. Hence, the achievement of pragmatic meaning is a matter of matching up the linguistic elements of the code with the schematic elements of the context. In addition to context, paralinguistic features are often considered during the analysis. Paralinguistic features include tone of voice, gestures, facial expressions, eye contacts of communicants that affect the communicative act and help decipher the implicit meanings and intentions of the utterances or statements that they make. Through the usage of relevant linguistic, pragmatic, paralinguistic, and extralinguistic means, the communicants establish the so-called *cooperative principle* that is the necessary precondition for mutual understanding and the realization of goals in the speech acts that they produce.

In summation, linguistic analysis is a complex and constantly developing process that involves multiple levels, procedures, and principles of scientific study.

Hybrid Disciplines

The following are the hybrid disciplines that have emerged as a result of merging linguistics with other disciplines in order to investigate issues of common interest. These disciplines make extensive use of linguistic frameworks: linguistic rules, concepts, and procedures.

Computational linguistics is the study of language using the techniques and concepts of computer science, especially with reference to the problems posed by the fields of machine translation, information retrieval, and artificial intelligence. Computational linguistics uses extensively syntactic rules and procedures for programming purposes; hence, the use of a linguistic framework is its indispensable part. In addition to this, computer programs are widely used to collect, systematize, and analyze vast corpora occurring in languages; that is, to produce and continuously extend corpus linguistics. Without these procedures, no level of modern linguistic analysis can be considered complete or final.

Neurolinguistics is the study of the neurological basis of language development and use in human

beings, especially of the brain's control over the processes of speech and understanding. Consequently, neurolinguistics constantly refers to semantic, syntactic, lexical, and pragmatic frameworks.

Psycholinguistics is the study of the relationship between linguistic behavior and the psychological processes that underlie it. The processes may involve memory attention, thought, and others. Hence, grammar, semantic, pragmatic, and discourse frameworks often lend to psycholinguistic studies.

Sociolinguistics studies the interactions between language and the structure and functioning of society. Pragmatic, discourse, lexical, and phonological frameworks are often used in sociolinguistic investigations.

Statistical linguistics studies the quantitative properties of language. Usage frequencies of acceptable or unacceptable linguistic structures are major objects of investigation. Therefore, the entire stock of linguistic frameworks is made use of at one point or another.

Mathematical linguistics studies mathematical properties of language, using concepts from fields such as algebra, computer science, and statistics, hence, wide usage of syntactic, phonological, and other linguistic frameworks.

Anthropological linguistics is the study of language variation and use in relation to the cultural patterns and beliefs of humankind, as investigated using the theories and methods of anthropology and referring to lexical, phonological, and discourse frameworks.

Applied linguistics is the application of linguistic theories, methods, and findings to the elucidation of language problems that have arisen in other domains. The term is especially used with reference to the field of foreign language learning and teaching, but it applies equally to several other fields, such as stylistics, lexicography, translation, and language planning, as well as to the clinical and educational fields. Since it covers a wide array of issues, it makes use of the entire linguistic framework.

Biological linguistics is the study of the biological conditions for language development and use in human beings, with reference both to the history of language in the human race and to child development. Syntactic, phonological, morphological, and lexical frameworks often lend themselves to biolinguistic studies.

Clinical linguistics is the application of linguistic theories and methods to the analysis of disorders of spoken, written, or signed language. Any linguistic framework may contribute to clinical studies depending on the topic of research.

Educational linguistics is the application of linguistic theories and methods to the study of the teaching and learning of a language, especially of a first language, in schools and other educational settings. Owing to its wide range of issues under investigation, the field of educational linguistics refers to any linguistic framework, depending on the circumstances of any particular study.

Ethnolinguistics is the study of language in relation to ethnic types and behavior, especially with reference to the way social interaction proceeds. Discourse, lexical, and pragmatic frameworks are of particular interest, but not exclusive to the field.

Geographical linguistics is the study of the regional distribution of languages and dialects, seen in relation to geographical factors in the environment. Phonological and lexical frameworks are particularly important to such studies.

Philosophical linguistics is the study of the role of language in the elucidation of philosophical concepts, and of the philosophical status of linguistic theories, methods, and observations. Pragmatic and discourse frameworks fall within the realm of these studies.

Theolinguistics is the study of the language used by scholars of biblical and other religious texts, by theologians, and others involved in the theory and practice of religious belief. The lexical framework lends itself to theolinguistic studies.

In conclusion, the overall field of contemporary linguistics itself provides an invaluable conceptual framework to a number of disciplines that cross-feed from its studies, findings, sources, concepts, and rules.

Mariam Orkodashvili

See also Artificial Intelligence; Information Theory; Linguistic Frameworks; Linguistics, Historical; Pragmatism; Semantics (Introduction to Theory); Semantics (Scientific and Empirical); Statistical Inference, Bayesian; Statistical Inference, Frequentist; Statistics; Structuralism (Introduction to Theory);

Syntax (Introduction to Theory); Syntax (Scientific and Empirical)

Further Readings

- Aarts, B. (2007). *Syntactic gradience. The nature of grammatical indeterminacy*. Oxford University Press.
- Aarts, B., David, D., Evelien, K., & Gergana, P. (Eds.). (2004). *Fuzzy grammar*. Oxford University Press.
- Biber, D., Stig J., Geoffrey, L., Susan, C., & Edward, F. (1999). *Longman grammar of spoken and written English*. Pearson Education Limited.
- Biber, D., Susan, C., & Randi, R. (2004). *Corpus linguistics. Investigating language structure and use*. Cambridge Approaches to Linguistics. Cambridge University Press.
- Boeckx, C. (2006). *Linguistic minimalism: Origins, concepts, methods, and aims*. Oxford University Press.
- Chomsky, N. (1957). *Syntactic structures*. Mouton Publishers.
- Chomsky, N. (1981). *Lectures on government and binding*. Foris Publications.
- Chomsky, N. (2002). *New horizons in the study of language and mind*. Cambridge University Press.
- Chomsky, N. (2007). *On language*. The New Press.
- Cook, G., & Barbara, S. (Eds.). (1997). *Principle and practice in applied linguistics*. Oxford University Press.
- Crystal, D. (1995). *The Cambridge encyclopedia of the English language*. Cambridge University Press.
- Crystal, D. (1997). *The Cambridge encyclopedia of language*. Cambridge University Press.
- Dancygier, B. (2006). *Conditionals and prediction. Time, knowledge and causation in conditional constructions*. Cambridge Studies in Linguistics. Cambridge University Press.
- Dancygier, B., & Eve, S. (2009). *Mental spaces in grammar. Conditional constructions*. Cambridge Studies in Linguistics, 108. Cambridge University Press.
- Haegeman, L. (2005). *The syntax of negation*. Cambridge Studies in Linguistics. Cambridge University Press.
- Halliday, M. A. K., & Ruqaiya, H. (1976). *Cohesion in English*. Routledge.
- Hart, C., & Dominick, L. (Eds.). (2007). *Cognitive linguistics in critical discourse analysis. Application and theory*. Cambridge Scholars Publishing.
- Heusinger, K. Von. (2002). Specificity and definiteness in sentence and discourse structure. *Journal of Semantics*, 19(3), 245–274.
- Hudson, R. (1999). *Sociolinguistics*. Cambridge University Press.
- Jackendoff, R. (2009). *Language, consciousness, culture. Essays on mental structure*. MIT Press.

- Jespersen, O. (1964). *Language. Its nature, development and origin*. Norton and Company.
- Johnson, K. (2008). *Quantitative methods in linguistics*. Blackwell Publishing.
- Kager, R. (1999). *Optimality theory*. Cambridge Textbooks in Linguistics. Cambridge University Press.
- Karttunen, L. (1976). Discourse referents. In J. McCawley (Ed.), *Syntax and semantics 7: Notes from the linguistic underground*. Academic Press.
- Labov, W. (2001). *Principles of linguistic change*. Blackwell.
- Lyons, J. (1972). *Introduction to theoretical linguistics*. Cambridge University Press.
- Meillet, A. (1948). *Linguistique historique et linguistique générale* [Historic and general linguistics]. Champion.
- Niyogi, P. (2006). *The computational nature of language learning and evolution*. MIT Press.
- Parikh, P. (2010). *Language and equilibrium*. MIT Press.
- Quirk, R., Sidney, G., Geoffrey, L., & Jan, S. (2000). *A comprehensive grammar of the English language*. Longman.
- Sapir, E. (1921). *Language*. Harcourt, Brace and World.
- Saussure, F. D. (1931) *Cours de linguistique générale* [Course in general linguistics]. Paris.
- Whorf, B. Lee. (1956). The relation of habitual thought and behavior to language. In J. B. Carroll (Ed.), *Language, thought, and reality: Selected writings of Benjamin Lee Whorf* (pp. 134–159). MIT Press.
- Widdowson, H. G. (1996). *Linguistics*. Oxford University Press.
- Woods, A., Fletcher, P., & Hughes, A. (1986). *Statistics in language studies*. Cambridge Textbooks in Linguistics. Cambridge University Press.

LINGUISTICS, HISTORICAL

Language change has been occurring as an aspect of human evolution throughout the centuries, serving as a sort of mirror reflection of human civilization and progress. The primary concern of historical linguistics is the study of language change. The central questions that historical linguistics raises are: Why do languages change? Why do certain languages die? How can we reconstruct extinct languages? How do borrowings from other languages affect the development of a language? Those questions will be addressed in what follows.

Such changes and development can be traced through all linguistic levels: phonetics, phonology,

morphology, syntax, and semantics. Moreover, as modern scientists claim, even discourse studies contribute significantly to the explorations of changes in the usage and pragmatic applications of certain grammatical and lexical constructions.

Different pronunciations of a word, different usages of words, the construction of new syntactic structures, the production of new linguistic units, and the disappearance of old versions all lend themselves to research in historical linguistics.

Linguistic change tends to occur when two equivalent forms coexist for a certain period and one form eventually gives way to another. As a consequence, one form disappears and another establishes itself in the language corpora. Historical factors such as territorial invasions, and hence language contacts, are often considered as driving forces of linguistic change. Additionally, with the development of trade, economy, technology, and intensification of social mobility, new lexical and phraseological units continually enter the language, new ways of pronouncing words are established, and new syntactic structures become fixed and more acceptable in any given language than the older ones.

Goals and Methods of Historical Linguistics

Among the primary goals of historical linguistics are to study the history of particular languages based on the existing written data; the study of prehistory of languages through comparative reconstruction; and the study of currently ongoing changes in the language that can be based on both written and spoken records and data. Through the study of these matters, historical linguists attempt to discover and formulate the general principles of linguistic change and identify the reasons for the transformations that languages undergo, both over centuries and at the time of study. Accordingly, historical linguists implement methodological principles of investigation that were proposed by Ferdinand de Saussure in the beginning of the 20th century and have been known as synchronic and diachronic linguistics since then. These methods of investigation, *synchronic* meaning observing changes in the present time and *diachronic* signifying language evolution through centuries, have been the fundamental principles in documenting linguistic

changes both historically and in the present time. Thus, it can be assumed that there is no single moment in which a living language is static and does not develop. The changes are ongoing, sometimes vividly and at others imperceptibly.

It should be noted that while the data of linguistic change from past centuries are relatively restricted because it is accessible only through written records, the data of present-day language development can be more diversified through written and spoken records. Modern historical linguists implement the various qualitative and quantitative methods of research and analysis, which may include interviews with informants, focus groups, participant observations, questionnaire analysis, online and live surveys, media analysis, and so forth.

It should also be borne in mind that it is not only the reconstruction of the past forms that is important for historical linguists, but interpretation, analysis, and cross-language comparison are of utmost significance for tracing certain regularities and patterns in linguistic changes.

Reconstructing a Proto-Language

One of the fundamental goals of historical linguistics is the reconstruction of proto-language that supposedly gave rise to the evolution of languages across the globe. Hence, the statement that the scientists make is that for all the differences that exist between languages, diverse languages might still originate from common proto-language. Therefore, the languages might be genetically related. Historical linguists present this statement most vividly through a family tree of languages that shows the connections of different languages to a common ancestor. Genetically related languages show common and recurrent correspondences in phonetic, lexical, or morphosyntactic structures. On the linguistic level, phonology and morphology seem to provide the most evident and illustrative data for historical analysis. Changes in sound systems and inflections are most frequently analyzed in the framework of historical linguistics to trace and establish genetic relations between the languages.

Reconstructing proto-Indo-European language, from which the majority of modern Euro-Asian languages derive, has been one of the most

fundamental landmark studies in historical linguistics conducted up to the present day. It should be noted at this point that such reconstruction does not lead to a primitive language. Reconstructed Indo-European language is a fully-fledged language, not a system of primordial grunts. A total of 8,000 years, or at most 10,000 years, is the furthest back that such type of reconstruction can take us to, and human language arose at least 80,000 years ago.

Several optimistically disposed scientists have gone as far as to reconstruct an even earlier language named *Nostratic* (Latin word meaning *our language*). According to the disputed views, *Nostratic* covers Indo-European; the Dravidian languages of India; the Kartvelian languages of the Caucasus; the Uralic family of languages, which includes Finnish and Hungarian; Altaic, which covers Turkish and Mongolian; and Afro-Asiatic, which includes Arabic and Berber. And some scientists have even taken a step further back and have put forward the idea of *Super-Nostratic* language, which links *Nostratic* to Arctic Indian languages in the north and to other African languages in the south. However, the idea of *Nostratic* and *Super-Nostratic* languages is controversial and requires further research and more valid evidence.

Typological Reconstruction

The most significant consequence brought about by the reconstruction method implemented by comparative historical linguistics is the elaboration of the typological reconstruction and especially the development of the subfield of typological syntax. The basic idea on which the method is based is the division and grouping of languages into basic types, each with its set of characteristics. English, for instance, can be categorized as verb-object (VO) language because it typically locates the object after verb. Depending on the syntactic structures predominant in a given language, it can be classified by language type; hence, the development of studies in typological syntax.

Typological reconstruction can also be applied to sounds, the most vivid example being the case of stop consonants [p, b, t, d] in Proto-Indo-European language, which are produced with a complete stoppage of airflow.

Thus, by historical reconstruction of a language, the detailed knowledge gained in the process of reconstruction can help the researchers to assign it to its corresponding type of languages. Identifying the type of language can provide further help in predicting the facts or characteristics for which there is no direct evidence.

Typological reconstruction is still a highly debated issue. There are no languages with fixed pure types. Most languages belong to a mixture of types owing to continuous fluctuations and alterations in the process of language development and evolution.

Tracing and analyzing the diffusion and spread of linguistic changes across different geographical areas is yet another newly evolving type of language reconstruction, which is still in the process of formation.

Language as a Constantly Changing Living Organism

One of the most fascinating ideas of language development and evolution was offered by Wilhelm von Humboldt (1767–1835), who perceived language as *energeia* (process, activity) and not *ergon* (end product of activity). Offering the deepest and most original insight into the nature of language, Humboldt thus presented language to humanity as a living and developing organism rather than a static compilation of forms and structures. He viewed language as an instant of speaking, an activity and not an inactive fixed form. Thus, Humboldt did not consider the study of a language's structure, its *Bau*, the ultimate end of linguistics. Because language in its fullest sense occurs only in the societal context in its acts of speech production and in what is being said through them, its true nature can only be explored and perceived in living discourse (*verbundener Rede*) and should be studied equally in its lasting manifestations in the works of culture and of science, in literature, poetry, and philosophy.

It is exactly this living nature of language that has permitted numerous scientists to conduct studies and explorations across languages and to come up with novel findings in language contacts and evolutionary processes, for it is through speech acts that linguistic forms tend to vary most

vividly and to constantly reintroduce novel meanings and forms.

Language Change and Its Causes

Language change is gradual and almost imperceptible. Historical linguistics mainly concerns itself with functional, psycholinguistic, and sociolinguistic explanations for language change. In reality, however, numerous other causes trigger language alterations over the course of centuries. The causes may range from random fluctuations in the usage of linguistic items to the natural and logical tendencies of humans to adapt linguistic usage and structures to evolving human speech apparatus and cognition.

Functional explanation views linguistic systems as self-regulating mechanisms that make languages more simpler, more symmetrical, and balanced. Phonological sound shifts, phonemic mergers, disappearance of homonymic words with conflicting meanings, and rectification of functional load (the number of functions a particular linguistic item has), all are subject to functional explanation.

Psycholinguistic explanation of change is concerned with cognitive processes going on in the human brain during the language acquisition process. Chomskyan principles of generative grammar that underlie the process of producing indefinite number of sentences using finite set of rules belong to the psycholinguistic explanation of change. The generation of novel types of syntactic structures and the disappearance of certain syntactic constructions could serve as illustrative examples in this respect.

The sociolinguistic explanation of change deals with the concepts of fuzziness and variation. *Fuzziness* implies the production of syntactic structures that are not quite grammatical, but are not totally ungrammatical either, meaning that individual speakers produce syntactic structures that do not strictly follow grammatical rules but that are not incorrect either. *Variation* implies the fact that speakers of a language have individual linguistic repertoires and use linguistic variables—different versions of linguistic items. The most illustrative examples could be speeches of different ethnic and social groups that produce linguistic items and structures idiosyncratic to them.

Alongside social variation, geographical variation of languages has been considered as one more contributing factor in linguistic change. It should also be mentioned that the linguist Edward Sapir noted in 1921 that linguistic drift (i.e., change) has a direction, meaning that the linguistic changes can be predicted and the direction of changes can be forecast. This directionality (mainly unidirectionality) principle has proved to be statistically significant and thus can serve as the basis for both linguistic evolution and language structure.

As a consequence of fuzziness and variation, certain linguistic items and structures disappear or become obsolete, while others spread widely and enter the corpora of a language. However, the concern that arises here is the basis upon which a construction will remain acceptable and spread widely and another one becomes regarded as unacceptable and disappears since there are no universally agreed restrictions and constraints on whether a certain linguistic item or construction is considered grammatically correct.

One noteworthy consequence that the study of language variations by historical linguistics has engendered is the establishment and wide use of the notion of linguistic variables—that is, sounds, morphological or syntactic elements, and lexical units that vary across times, social strata, and geographical areas, and that have been the objects of investigation in the field of statistical linguistics for a long time.

Sound Change

Probably the best-researched realm of historical linguistics is phonological change, which may take place both on the level of phonemes (phonemic change) and on the level of speech production (phonetic change). Sound change takes place when the ways of production of vowels or consonants change under the influence of the neighboring sounds. The basic assumption of comparative historical linguistics is that sound changes are consistent or regular rather than haphazard. Furthermore, if there are consistent sound correspondences between words with similar meanings in languages where borrowing can be ruled out, the correspondences cannot be due to chance. It can be, therefore, inferred that the languages concerned are so-called daughter languages descended from one

parent language. To illustrate the point, the following correspondence between English and Latin could provide a good example. Where English has *f*, corresponding Latin words with similar meanings have *p*, such as *father-pater*, *foot-pedem*, *fish-pisces*. When Romans arrived in Britain, the natives did not yet speak English, which was introduced at a later stage, so borrowing from Latin into English can be excluded. Consequently, the explanation of correspondences can lie in the fact that both Latin and English derived from one parent (proto) language that existed earlier.

The following sound change patterns can be identified: *palatalization* (fronting of the raised part of the tongue, e.g., mouse–mice, foot–feet); *velarization* (the movement of tongue backward toward velum or soft palate, e.g., English *sword* and German *Schwert*); *diphthongization* (changing of a simple sound into diphthong, e.g., house, ride from Middle English [u:] and [i:]); *monophthongization* or leveling of diphthongs before [ə] as in fire, tower, resulting in [fa:], [ta:]. Consonants can also change the manner and place of their articulation. The most common change pattern is *spirantization*, the change from stop to fricative (spirant), as was the case in Grimm's Law, when the Indo-European voiceless stops [p, t, k] became voiceless fricatives [f, θ, x] in the Germanic languages.

Phonetic change is most frequently conditioned by the specific phonetic environment that it occurs in. As an instance, the sound [r] was lost before consonants and word-finally—that is, court [ko:t], hair [heə]—but was retained in other positions—for example, hairy [heəri], right [rait]. Vocalization, disappearance, or weakening of consonants are yet further instances of sound changes.

The most widely spread type of changes that sound modifications lead to is *assimilation*, the partial or total adjustment or resemblance of a sound to another one, often the neighboring one, for example, im-possible, il-literate [m-p; l-l]. The word *assimilation* itself illustrates the point most vividly. It derives from the Latin word *as-similare* that through the assimilation process was derived from the earlier version *ad-similis* (meaning *like*).

A rarer case of sound change is *dissimilation*, whereby the neighboring sounds become partially or totally different, for example, pilgrim, Latin *peregrinus*.

Further instances of sound change are chain shifts, as in Grimm's Law, where the most illustrative examples are the change of Indo-European /p, t, k/ into /f, θ, x/, and /b, d, g/ into /p, t, k/. Another famous instance of sound change is the *Great Vowel Shift*, which involved the raising of long vowels, for example, /e:/ into /i:/, as in *meet*.

Grammatical Change

Grammatical change involves morphological and syntactic changes. As the second most frequent type of change after phonological change, the morphological change has facilitated the typological grouping of languages according to inflectional morphemic features. Hence, there exist isolating, agglutinating, and inflecting types of languages. In isolating languages (e.g., Chinese), words consist of single morphemes. In agglutinating languages, words consist of more than one morpheme, but each morpheme is formally separable and has a single meaning (e.g., Japanese). In inflecting languages, morpheme boundaries are often blurred, and several meanings can be accommodated in one form (e.g., Greek and Latin). In the process of linguistic change, languages can turn from isolating into agglutinating, and thereafter into inflectional. The cycle may as well repeat several times. Changes from isolating structures to agglutinating structures often happen in pidgins and creoles, where unstressed words often become prefixes and suffixes as a consequence of phonological reduction.

Modern English has lost most of its inflections and has turned from an inflecting synthetic language into an analytic language, in which fixed word order has substituted the former functions of inflections.

The most illustrative example of syntactic change is the development of tense forms through centuries. For instance, in Old English, there were only two tenses, present and past. Other tense forms developed later, most conspicuous opposition being simple versus progressive forms. In addition, questions and negatives were formed without the auxiliary verb *do* in Old English, for example, *I know not, whom trust you more?*

The types of syntactic changes fall under two major categories, reanalysis and grammaticalization, *reanalysis* implying the changes in word

order patterns in sentences (e.g., the development of SVO structure in English), and *grammaticalization* (or *grammaticization*), illustrating the development of tense categories (e.g., progressive forms), the development of main verb forms from auxiliary ones (e.g., the verb *will* acquiring the meaning of *want*), or the development of negative particles from negative intensifiers (e.g., not . . . a step). The term *grammaticalization* or *grammaticization* was coined by a French linguist, Antoine Meillet, who defined it as the attribution of a grammatical character to a previously autonomous word. The most illustrative case in this respect is the fact that verbs of *volition*, such as *want*, *wish*, and *desire*, have become the future markers; that is, have acquired the grammatical category of expressing the future.

Vocabulary Change

Several linguistic processes can be reflected in vocabulary change. These may include new word formation, word coinages, compounding (constructing one word from two independent words), affixation (adding prefixes and suffixes to a word to coin a new lexical unit), narrowing and extension of meaning (when a word loses one of its meanings or acquires a new meaning), amelioration (improvement of meaning, e.g., *knight* in Old English meant "boy, attendant") or pejoration (a negative evaluation of meaning, e.g., *silly* in Old English meant "happy, blessed"), production of new metaphors, metonymies, and so forth. Languages coming into contact with each other most frequently contribute to vocabulary changes. New lexical or phraseological units are often borrowed from one language into another. The new words often fuse into the host language through affixation or derivation that puts the loan words in the shape acceptable for the host language.

Borrowing a new word from another language that has a corresponding word in the native language may lead to the disappearance of one of those. Hence, the transformation of corpora takes place.

The reasons for vocabulary change in a language are various, meeting new communicative requirements being the most frequent ones. Furthermore, a psychological factor in vocabulary change is the attempt to avoid using taboo words and not to make direct reference to the unpleasant

or stigmatizing concepts and words. In addition, if the word usage wears off—if it is widely used and loses its expressive connotation—a new, more emphatic word enters the corpora and replaces it. Further instances of vocabulary change are related to the developments in science, technology, especially in fast growing and changing information technologies. The growth of trade, tourism, and of transnational business enterprises makes the process of vocabulary borrowings and linguistic interference even faster.

Relations of Historical Linguistics to Other Discipline

Historical linguistics has always been cross-feeding from other disciplines, such as geology, biology, statistics, and sociology. As early as the 19th century, the uniformitarian principle entered the realm of historical linguistics. Genetic linguistic relationships have been long displayed through the family tree model provided by biology. Furthermore, evolutionary biology has triggered discussions on language development and change and has consequently given rise to the new field of evolutionary linguistics, for which evolutionary theory has been the focal point. It perceives linguistic evolution as a progression from a more primitive to an ideally developed state of language that will be followed by decay and finally, lead to language death. Another view, however, puts linguistic evolution within the frame of natural selection, where random mutation leads to change. The role of genes and memes in replication and copying processes has also been adopted by evolutionary linguistics to trace the production and replication or decay of certain linguistic items and structures through time. Another important concept adapted from biology is the notion of *gradation*, which helps historical linguists arrange the linguistic forms and their changes in graded order (meaning not drastic or extreme alterations but changes organized in graded sequence). The aforementioned notion of *fuzziness* could serve as an example, implying that there are not totally incorrect or absolutely correct linguistic structures; rather, linguistic evolution constantly produces partially correct and partially incorrect items.

Mariam Orkodashvili

See also Linguistics, Contemporary; Semantics (Introduction to Theory); Semantics (Scientific and Empirical); Syntax (Introduction to Theory); Syntax (Scientific and Empirical)

Further Readings

- Aarts, B., David, D., Evelien, K., & Gergana, P. (Eds.). (2004). *Fuzzy grammar*. Oxford University Press.
- Aitchison, J. (1991). *Language change: Progress or decay?* Cambridge University Press.
- Crystal, D. (2002). *Language death*. Cambridge University Press.
- Gamkrelidze, T., & Vjaceslav I. (1995). *Indo-European and the Indo-Europeans: A reconstructions and historical analysis of a proto-language and a proto-culture* (Trends in Linguistics: Studies and Monographs 80, 2 vol. set). Mouton de Gruyter.
- Hudson, R. (1999). *Sociolinguistics*. Cambridge University Press.
- Humboldt, W. von. (1836). *Ueber die Verschiedenheit des menschlichen Sprachbaus und seinen Einfluss auf die geistige Entwicklung des Menschengeschlechts* [The diversity of human language-structure and its influence on the intellectual and spiritual development of mankind]. F. Dümmler.
- Labov, W. (2001). *Principles of linguistic change*. Blackwell.
- Meillet, A. (1948). *Linguistique historique et linguistique générale* [Historic and general linguistics]. Champion.
- Sapir, E. (1921). *Language*. Harcourt, Brace and World.
- Saussure, F. de. (1931). *Cours de linguistique générale* [Course in general linguistics]. Otto Harrassowitz Verlag.
- Schendl, H. (2001). *Historical linguistics*. Oxford University Press.
- Steponavičius, A. (1987). *English historical phonology*. High School Publishing.
- Woods, A., Paul, F., & Arthur, H. (1986). *Statistics in language studies*. Cambridge Textbooks in Linguistics. Cambridge University Press.

LOGIC, FORMAL AND INFORMAL

Mathematical logic aims at a universalist characterization of kinds of structural relationships that may be studied nonempirically. It thereby offers a resource for all fields, including the study of cognitive structures and semantic structures that constitute reasoning well. According to the logicist thesis espoused in Alfred North Whitehead

and Bertrand Russell's *Principia Mathematica*, mathematical logic is a field of its own—a field to which the various branches of pure mathematics, including geometry, all belong. Whether this is correct or whether mathematical logic is itself a branch of mathematics is an ongoing question much debated. In any case, it is clear that many different structures may be studied, and among them are what are called *deductive logics* (e.g., classical, intensional, modal subjunctive [counterfactual], intuitionist, relevant, paraconsistent), each offering axioms and/or inference rules of transition (based on syntactic *formal* features alone) meant to capture all and only those well-formed formulas (WFFs) that are *valid* (i.e., preserve a fixed distinguished value under all *admissible* interpretations).

Informal logic endeavors to explain and teach real-life argument practices that are regarded as *good* by the standards that are used as paradigms in various fields. The notion has been understood and interpreted in a variety of different ways, and its connections with formal logic are often debated. Over the years, diverse disciplines have endeavored to provide it a proper home under each discipline's conception of what constitutes reasoning well in general—using its own field more or less as a paradigm. The presentation of informal logic may take different forms as it becomes tailored to applied disciplines such as cognitive (and developmental) psychology, evolutionary anthropology, and law. Various textbooks offer many pedagogical innovations drawn from a wide array of fields such as applied probability and statistics, economics, cognitive psychology, linguistic semiotics, speech communication, and rhetoric; and may also include doctrines of historicism and historical materialism. Considered in this way, the field of informal logic is a multi-disciplinary attempt to develop an argumentation theory that endeavors to give an account of what constitutes real-life good reasoning. This entry first explores the traditions within which logic itself has been defined and then continues with an extended discussion of the distinctions between the nature and methods of argument of formal and informal logic, including explanations and examples of the various approaches, analytical methods, and arguments of rival schools up to the present day.

One Hundred Definitions

The philosopher, logician, and mathematician Charles Sanders Peirce (1839–1914) amusingly suggested that nearly one hundred definitions of *logic* may have been tried, but he thought it had been generally conceded that logic presents the methods and principles that distinguish good from a bad argument (reasoning). The Scottish philosopher David Hume (1711–1776) had famously contended that reason is the slave of the passions—that is, a tool for successful means–end calculations. Its value lies, thereby, in its success in guiding one to achieve one's goals, and good reasoning is the best avenue. But this view of the nature of reason remains controversial, and Immanuel Kant (1724–1804) famously held, contra Hume, that it is reason that sets goals and determines worthy values.

When reason is viewed as an evolved set of cognitive faculties for engagement, logic is viewed as being within the province of (evolutionary) psychology. This is the pragmatist view of John Dewey (1859–1952), for whom the field of *logic* (whether formal or informal) endeavors to discover and improve the behavior patterns of human cognitive inference that have evolved because of the adaptive advantages they conferred on the species. Reasoning poorly, especially the common mistakes in applied probability, is thereby diagnosed in terms of behaviors that are a hold-over from the natural evolutionary history of the brain.

William James (1842–1910), who with Peirce originated philosophical *pragmatism*, put no emphasis at all on viewing the methods of good argumentation as something abstractly tracking truth and fact. This is usually regarded as its Achilles heel. Obviously, the most fallacious arguments may, sadly, enjoy an interval of practical success—especially if *success* can be measured with respect to a community caught up in the dogmatism of an age. A careful pragmatist may address this by defining “good reasoning” as the ideal kind of reasoning that results from an *unencumbered* practice of trial and error. Defining *idealization*, however, will require the mathematical logic of *limits*, and such abstract mathematical notions remain difficult for pragmatism and evolutionary psychology to embrace. Moreover, success at the limit of idealized practical engagement still cannot secure truth. The pragmatist responds by offering a notion of *truth* as whatever it does is secure.

Peirce and James wrote before (or near) the dawn of mathematical logic, which altered the debate significantly, subsuming at least part of mathematics itself into the domain of its new conception of *logic*. However, there remain rival traditions within mathematical logic itself. On the algebraic tradition, deductive logic began with Aristotelian categorical forms:

A: All S are P E: No S are P

I: Some S are P O: Some S are not P

It studied the syllogism and offered rules governing *immediate inference* such as contraposition (switch S and P and complement each), obversion (change quality, universal $\Leftrightarrow$ particular, and add a complement to P), and inversion (switch S and P). Subsequent medieval scholastic studies introduced propositional connectives, and endeavored to distill quantifier scope to block fallacious inferences from “I owe you a donkey” to “some donkey is such that I owe it to you.” The invalidity of a syllogism was thought to consist in it having a counterexample with substitutions for S, P, and middle term M. Thus consider:

All P are M; All S are M; Therefore, Some S are P

This is shown to be invalid by appeal to:

All cats are animals; All dogs are animals
Therefore, Some dogs are cats

The accepted counterexample makes the premises true and conclusion false.

Algebraic approaches far transcend what an Aristotelian can accomplish with categoricals. It began with the work of George Boole (1815–1864) and with Peirce and Ernst Schröder’s *Algebra of Logic*, which endeavored to combine algebras of operations to which *and*, *or*, and *not* can be associated (product, sum, inverse) with quantifiers understood as (possibly infinite) product and sum. Booleans and set theorists removed the Aristotelian restriction that A, E, I, O forms apply only when the categories are not empty. Modern-day studies in formal logic emphasize the algebraic tradition, focusing on valuation tables (i.e., truth-tables) for determining validity (understood as preservation of a distinguished value) and quantification theory with identity. The algebraic tradition presents an algebra of fixed operations. Validity no longer

relies on an appeal to substitution counterexamples but is determined by considering algebraic rules or Venn diagrams picturing the premises without the conclusion. This begins the separation of the field of logic from anything empirical. It frees logic from all considerations of the psychology of inference and the pragmatics of success. But it accepts the Humean barb that in any deduction (as opposed to induction) the conclusion cannot bring forth any content not already in the premises.

The rival tradition in mathematical logic originates with Gottlob Frege (1848–1925), who rejects the algebraic approach with its fixed operations. Frege made the study of operations merely one of a great many applications of his logic of functions, which investigates *a priori* (without empirical investigation of nature) structural features of relations between functions but also assures the existence of ever new functions themselves. In Frege’s logic, one discovers the fact that every relation has a converse, but that the converse of a function may not itself be a function. (The trigonometric *sine* of an angle is a well-known example.) This discovery, abstract as it is, should guide one’s practical inference—if good practical inference aims at *truth*. You may love but one, but you’ll behave ignorantly if you conclude that your beloved cannot be loved by a rival. It remains improper as a conclusion even if it happens (by some sheer contingency) to be successful. No feature of psychology, evolution, social practice, history, and so on, is the least bit relevant.

A Fregean nonempirical tradition in mathematical logic allows for the discovery of ever new functions, and thereby the Humean barb is rejected. A deduction’s conclusion may well contain new information not obtainable by the algebraic process of manipulating the operations given in the premises.

The tradition finds the notions of *number*, *successor*, *infinite*, and *finite* to be purely quantificational notions concerning relations. Fregean logic assures the existence of relations involved, for example, in the assertion that for any one-to-one relation *R* whose relata are correlated with a group of exactly 5, there are groups of exactly 2 and 3 and a relation *S* isomorphic to *R*. (For example, digits on my left hand are correlated one-to-one and onto the digits of my right hand—namely, the group of thumb and forefinger, and the group of middle, ring, and pinkie.) There is no known bridge between the Fregean and algebraic

traditions *within* mathematical logic, and quite different philosophies of mind, rationality, and reason have grown up around each—with those connected with the Fregean more rationalist and those surrounding the algebraic distinctly empiricist.

Deductive Logic (Informal vs. Formal)

Informal logic has the mission to “study real-life argument.” This can take many different forms and involve eclectic methods borrowed from several fields. The general methods of argument analysis explore aspects of burden of proof, norms governing style (in specific fields such as law, politics, empirical scientific research) and may include historical, social, libertarian, neoliberal, antiliberal, feminist, and historical materialist points of view. It points out complex assumptions (philosophical, epistemological, social-political, etc.) that may lurk behind premises but are not explicitly mentioned in premises of any real-life argument.

What, then, is the practical value of formalism? In deductive formal logic, one works with the formalisms themselves, and the rules of transition are based solely on the formal syntactic features. Formal logic may draw attention to the challenging practice of transcription from natural language to formal notations. In this respect, the formal logician is not disagreeing with the informalist. The disagreement arises because the formal logician holds that transcription is *required* for the evaluation of a deductive argument’s being *valid*—since *validity* in (for example) quantification theory is understood as a feature arising from the structure that is given by *and*, *or*, *not*, *all*, and *some*. (Any admissible model of its premises thereby models its conclusion.) To illustrate, consider:

This argument is valid; Therefore, the cat is on the mat and the cat is not on the mat.

According to the formalist, in order to assess the validity of the above argument, one must first transcribe it to determine its logical form. If one doesn’t do this, a danger looms. One might find no counterexample, wherein the premise is true and the conclusion false. One might erroneously conclude that the argument is valid! But then its contradictory conclusion should be true; we get a paradox. The formalist’s demand to know the

structure of the argument diagnoses the mistake. Suppose it is this:

q ; Therefore, p and $\sim p$

The formal logician concludes that the deductive argument is simply invalid.

The transition (inference) rules governing $\bullet$, $\vee$, $\sim$ (*and*, *or*, *not*) are the foundation of the De Morgan’s rules governing the operations $\cap$, $\cup$, $\sim$ (*intersection*, *union*, and *complement*). Using $\Leftrightarrow$ for transition:

$$S \cap P \Leftrightarrow S \cup P \quad \sim(p \bullet q) \Leftrightarrow \sim p \vee \sim q$$

$$S \cup P \Leftrightarrow S \cap P \quad \sim(p \vee q) \Leftrightarrow \sim p \bullet \sim q$$

The ideas are given by our natural understanding of *and*, *or*, *not*, and it is these that give rise to the operations. Consider also quantifier negations:

Not all S are $P \Leftrightarrow$ Some S are not P

$$\sim(x)(\sim Sx \vee Px) \Leftrightarrow (\exists x) \sim(\sim Sx \vee Px)$$

No (i.e., not some) S are $P \Leftrightarrow$ All S are not P

$$\sim(\exists x)(Sx \bullet Px) \Leftrightarrow (x) \sim(Sx \bullet Px)$$

The quantifier and De Morgan’s transition patterns are natural.

For quantification, the mathematics of *sum* and *product* of the Peirce–Schröder *Algebra of Logic* proved tedious and clumsy in practice, and the Whitehead–Russell (W–R) notations became widely adopted instead. We now have:

Aristotle	Set (Venn)	W–R
A: All S are P	$S \cap P = \emptyset$	$(x)(\sim Sx \vee Px)$
E: No S are P	$S \cap P = \emptyset$	$\sim(\exists x)(Sx \bullet Px)$ i.e., $(x) \sim(Sx \bullet Px)$ i.e., $(x)(\sim Sx \vee \sim Px)$
I: Some S are P	$S \cap P = \emptyset$	$(\exists x)(Sx \bullet Px)$
O: Some S are not P	$S \cap P = \emptyset$	$(\exists x)(Sx \bullet \sim Px)$ i.e., $(x) \sim(Sx \bullet \sim Px)$ i.e., $(x)(\sim Sx \vee \sim \sim Px)$ i.e., $(x)(\sim Sx \vee Px)$

The power of the W–R transcription is nicely illustrated in its ability to easily capture the following:

All A are B ; Therefore, all heads of some A are things that are heads of some B .

$$(x)(\sim Ax \vee Bx); (x)(\sim (\exists y)(Ax \bullet Hxy) \\ \vee (\exists y)(Ax \bullet Hxy)) \cdot$$

No categorical treatment of this argument is workable.

The great benefit of formalism is that no ambiguities can arise. Real-life negations, predications, quantifications, conjunctions, and disjunctions can be very difficult because they are often expressed in many different and often misleading ways in natural language. A colloquial way to say

“Not all snakes are poisonous.”
is to say, with a tone of voice:

“All snakes are not poisonous.”

One must be on the lookout for such poetic (if you will) devices. Informal logic addresses itself to this. The study of real-life natural language modes of expression is very important, but not a rival to formal logic emphasis on structure. Indeed, it is guided by it.

Consider the following formal pattern of the immediate inference that is *Contraposition* (the second of which is parameterized):

All *S* are *P* $\hookrightarrow$ All *P* are *S*

All *A* and *S* are *A* and *P* $\hookrightarrow$ All *A* and *P* are *A*
and *S*.

Parameterized immediate inference is what often occurs in natural language cases. We don't use:

All life forms indigenous to the Andromeda galaxy are silicon-based life forms.

Therefore, All non-silicon-based-life-forms are
non-life-forms indigenous-to-the-Andromeda-galaxy.

The proper contraposition is:

All life forms not silicon based are life forms not
indigenous to the Andromeda galaxy.

Knowing the formal pattern is quite illuminating. Natural language complements can be challenging. Because of the many styles and meanings of negation, false complements abound. For example, *not believe* does not mean *disbelieve*, *inhuman* does not mean *not human*, *not proved* does not

mean *disproved*, and *unhappy* does not mean *not happy*; whereas *unintelligible* does mean *not intelligible*. Note that *not valid* does not mean *invalid*. *Valid* and *proof* are technical notions of formal deductive logic. Only a deductive argument that is not valid is properly said to be *invalid*. Every inductive argument fails to be valid. That is not a fault since it wasn't purporting to be a deduction. Empirical theories such as evolution by natural selection are not proved, but that is not a fault, because proof is not sought and belief is warranted by a wealth of corroborating evidence in their favor.

Informal logic often focuses on fallacies of reasoning (both formal and informal). Some schools of informal logic go so far as to argue that there are successful cognitive operations that are the victors of the processes of natural selection. In this school, controversies over whether such cognitive operations might be successful and yet widely depart from truth is not the focus of attention. Instead, it is the norms of acceptance within a field that is the focus. But other schools of informal logic do not agree that paradigms of good reasoning in the various fields are the proper arbitrators. Such schools endeavor to preserve the thesis that *poor reasoning* is distinctive in its failure to capture truth and fact.

Natural language makes it challenging to parse negations and double negations. Formal characterization eliminates any chance of ambiguity. This suggests that the study of formal and informal fallacies may be best understood as simply the study of the myriad ways of confusion due to equivocations arising in natural language. Whether it is applied deductive logic, applied inductive methods, or even applied mathematics and physics, the kinds of equivocations and confusion that arise due to ambiguities of natural language uses of *and*, *or*, *not all*, and *some* can be enormous. We have the ever-popular example: “Jack and Jill went up the hill.” But this is not the *and* of conjunction. They went up the hill together in consort with one another. It does not mean Jack went up the hill *and* Jill went up the hill.

Consider the case of the formal fallacy of *Affirming the Consequent*. It is simply a problem of confusion due to equivocation. Consider the Disjunctive Argument patterns:

Either *p* or *q*; not *p*; Therefore *q*.

Either not *p* or *q*; *p*; Therefore *q*.

The letters p and q stand in for declarative statements. Admittedly, by making mistakes about whether one has a genuine double negation in the natural language, one might misapply the disjunctive argument patterns. That is a case of equivocation. Now the expression “Either not p or q ” can be clumsy, and logic books often stipulate that “If p then q ” will conveniently be used as a substitute. With the stipulation, we get:

If p then q ; p ; Therefore q .
i.e., Either not p or q ; p ; Therefore q .

It is difficult to imagine someone reasoning with:

Either not p or q , q therefore p .
i.e., If p then q ; q ; Therefore, p .

However, if equivocation is at play, one might well reason with:

Events of kind q are (very typically) caused by events of kind p ; q ; Therefore (very likely) p

This reasoning is inductive. The mistake lies in conflating the natural language “if ... then,” which may express causal ties, with the stipulated “if ... then” of the logic books.

Ambiguities also infest the use of “if ... then” because it may convey a subjunctive. We have:

If p were then q would be. $p \Box q$.

Consider the following:
 No student would have passed were they to goof off.
 $\sim(\exists x)(Sx \bullet (Gx \Box Px))$

Compare this:
 Every student would have failed were they to have goofed off.
 $(x)(\sim Sx \vee \sim(Gx \Box Px))$

The formalisms are logically equivalent, but the English of the second is ambiguous and might mean:

$(x)(\sim Sx \vee (Gx \Box \sim Px))$.

Equivocating, one may be misled into thinking we have:

$(x)(\sim Sx \vee \sim(Gx \Box Px)) \not\equiv (x)(\sim Sx \vee (Gx \Box \sim Px))$.

Obviously, equivocation does not support belief in the law and inference of conditional excluded middle:

$(p \Box q) \vee (p \Box \sim q)$
 $\sim(p \Box q) \not\supset (p \Box \sim q)$

The formalisms make the error clear, where the English struggles. Thus are some of the benefits of formal logic working in collusion with informal logic.

Are there really *Formal Fallacies* that are committed by untrained reasoners or just various ways of equivocation? Consider the formal fallacy of *Begging the Question* whereby the conclusion is asserted in one or another premise. A good case is:

Action a is a murder; Therefore action a is wrong.

By definition, “ a is a murder” means “ a is wrongful and a is a killing.” Thus, the conclusion is asserted in the premise. It is only equivocating on the meaning of *murder* that one might fallaciously think that the argument was informative. Note that the premise, in this case, asserts a conjunction, which by its nature asserts each conjunct. The following is not a case of begging the question:

Everything is physical; e is an electron; Therefore, e is an electron and e is physical.

The conclusion is not asserted by the premise but is, instead, such that any model of the premise models the conclusion. Any instance of the deductively valid argument form

p ; Therefore p

is an argument that begs the question. But the fallacy of thinking such an argument to be informative seems to arise only via equivocation.

The *Gambler's Fallacy* is another case. Each time one flips an idealized fair coin the mathematical probability of it being Heads is $\frac{1}{2}$. But what is the probability, after, say, three flips, none being Heads, of the fourth flip being Heads? The answer still is $\frac{1}{2}$, at least on one interpretation of the question, for the current flip event is not impacted by the past. The gambler, however, may remark: “But surely Heads is more likely *now*.” What is at work in the gambler's misguided thinking? There is an ambiguity in the meaning of *now* as used

here. What is the probability in four flips of a fair coin (which is the same as one flip of four fair coins) that *none* should be Heads? Answer 1/16. The fallacy here is equivocating on the import of *now*, and thereby conflating two different events with distinct probabilities.

Perhaps a most intriguing case of a formal fallacy is illustrated by the *Conjunction Fallacy* (Linda Paradox), of Amos Tversky and Daniel Kahneman, which allegedly infests ordinary uses of the notion of *probability*.

In college (in the 1980s during the heyday of the feminist movement), the woman Linda was deeply concerned with such issues of discrimination and social justice, and she participated in demonstrations. Which is more probable:

- (a) Linda is a bank teller
- (b) Linda is a bank teller and is active in the feminist movement.

In empirical studies, most participants chose (b) in striking violation of the mathematical probability law that $\text{Prob}(A \mid B) < \text{Prob}(A)$. That anyone would commit such a fallacy seems shocking. Researchers have disagreed about whether the explanation concerns an evolved cognitive weakness in thinking about probability by thinking of *representative* cases based on parochial experience—that is, of women of the day—rather than applying abstract mathematical or frequency or physical propensity notions of probability. But at least one school, as made clear in Hertwig and Gigerenzer (1999), faults the experimental methods, noting ambiguities in the question's framing and wording. Perhaps (a) was interpreted by participants to mean:

- (a)*: Linda is a bank teller and *not* active in the feminist movement.

If so, then most participants got it right. Research in informal logic reasoning must endeavor to avoid infestation of such experiments by equivocation and confusion.

Equivocation is rightfully salient among the informal fallacies that informal logic draws special attention to. It may well also be at the heart of the common *Informal Fallacies*? Typical categories of fallacies given in informal logic include the following: *Fallacies of Relevance* (pity, ad populum,

ad hominem, accident, straw person, non-sequitur, red herring); *Fallacies of Weak Induction* (unqualified authority, ignorance, false cause, hasty generalization, slippery slope); *Presumption Fallacies* (begging the question, complex question, false alternatives); *Fallacies of Ambiguity* (equivocation, amphiboly); and *Whole-Part Fallacies* (composition, division). All admit that it is not clear whether a perfectly justified classification scheme can be found.

Whole-Part Fallacies notoriously equivocate on the wide variety of meaning of *part*. A member is not a part of a class in the same sense that a piston is part of an engine, a governor is part of the State government, a chromosome is part of a cell, a finger is a part of a hand, and a galaxy is part of the universe. Consider the famous informal fallacy Ad Hominem, whereby one allegedly evaluates an argument by criticizing the character (and/or associations) of the person giving it. Perhaps what really happens is triage—deciding inductively that if a person has a sordid history of giving (or belongs to an association that gives) bad arguments intended to deceive, the current one (on a relevantly similar topic) is likely also to be bad. Unjustified triage ought to be avoided, but it is not a fallacy. Is there a fallacy of Appeal to Pity whereby one questionably evaluates an argument to a conclusion (say guilt or nonguilt of a defendant) simply by one's being justified in having pity over the miserable state of those concerned? Appeal to the pitiable state might well give good grounds for practical action (e.g., lighter sentence for the convicted), but that simply changed the issue entirely. Equivocation is again at work here as it is in the fallacy of Appeal to Force, whereby the notion of a *reason to believe* can mean expedience to bring oneself to believe (say, to stay alive) as opposed to meaning to bring about a belief *based on the evidence*. It begins to look as though a great many of the so-called informal fallacies are ways of equivocation. If that is the case, it is a very worthy topic of study for informal logic and it requires on-the-ground, real-life investigations to understand the many ways.

Informal logic should not be confused with *inductive logic*, which studies such topics as statistics, expected utility, Bayesian theory, and probability (mathematical, causal/propensity, frequency, degree of credence). Statistic and probability theory

are applied mathematics. Bayes' theorem, after all, is a theorem of pure mathematics. Applying it, however, can be quite a complicated matter of discerning whether there is genuine independence of events, and properly recognizing how orderings are required for infinitary as opposed to finitary probability spaces. All the same, applications require transcriptions into formal mathematical apparatus, and it is here that confusions produced by natural language often arise. Informal logic has just as significant a role to play when it comes to errors that can arise in transcriptions involving applied mathematics, probability theory, applied statistics, and so on—even applied physics. Its role here, however, seems no different from its legitimate role in the study of how natural language ambiguities may infest applications of formal theories.

Gregory Landini

See also Deduction; Hypothetico-Deductivism; Inference; Logic, Inductive; Logical Theory

Further Readings

- Battersby, M., & Sharon, B. (2011). Fallacy identification in a dialectical approach to teaching critical thinking. *OSSA Conference Archive*. University of Windsor.
- Bowell, T., & Gary, K. (2014). *Critical thinking: A concise guide* (4th ed.). London.
- Browne, M. N., & Stuart, M. K. (2018). *Asking the right questions: A guide to critical thinking* (12th ed.). Prentice Hall.
- Cocchiarella, N. B. (2007). *Formal ontology and conceptual realism*. Springer.
- Copi, I. (1979). *Symbolic logic*. MacMillan.
- Eemeren, F. H. van, & Rob, G. (1992). *Argumentation, communication, and fallacies: A pragma-dialectical perspective*. Lawrence Erlbaum Associates.
- Hurley, P. J. (2008). *A concise introduction to logic* (10th ed.). Thomson Wadsworth.
- Govier, T., & Ayers, L. (2012). Logic, parables, and argument. *Informal Logic*, 32(2), 161–189.
- Groarke, L. A., Christopher, W., & Tindale, L. C. (2021). *Good reasoning matters!* (6th ed.). Oxford University Press.
- Hatcher, W. S. (2014). *The logical foundations of mathematics*. Pergamon.
- Hertwig, R., & Gigerenzer, G. (1999). The 'conjunction fallacy' revisited: How intelligent inferences look like reasoning errors. *Journal of Behavioral Decision Making*, 12(4), 275–305.
- Hughes, W., Jonathan, L., & Katheryn, D. (2014). *Critical thinking: An introduction to the basic skills* (7th ed.). Broadview.
- Kahane, H., & Nancy, M. C. (2002). *Logic and contemporary rhetoric: The use of reason in everyday life* (9th ed.). Wadsworth.
- Lewis, D. (1973). *Counterfactuals*. Blackwell.
- Puppo, F. (2019). Informal logic: A 'Canadian' approach to argument. *Windsor Studies in Argumentation*, 9.
- Seay, G., & Susana, N. (2012). *How to think logically* (2nd ed.). Pearson.
- Stalnaker, R. (1968). A theory of conditionals. In R. Stalnaker (Ed.), *Studies in logical theory*. Blackwell.
- Tversky, A., & Kahneman, D. (1983). Extension versus intuitive reasoning: The conjunction fallacy in probability judgment. *Psychological Review*, 90(4), 293–315.
- Weston, A. (2018). *A rulebook for arguments* (5th ed.). Hackett.
- Wilson, D. C. (2020). *A guide to good reasoning: Cultivating intellectual virtues* (2nd ed.). University of Minnesota Libraries.

LOGIC, INDUCTIVE

Logical arguments are generally categorized as either inductive or deductive. Induction can be defined as a process for moving from particular instances to general statements or conclusions. It starts with a number of singular statements based on specific instances of an event. Based on these statements, a general conclusion is made. It is a method that helps us move from particular facts or observations to general statements that comprehend them. It is, therefore, used to generate theory from data. It is often seen as the opposite of deduction. This entry describes the nature of inductive logic and provides a historical account of its origins in ancient Greece and its subsequent development. Next, the major critiques of the inductive method are presented, followed by responses from its supporters, and the entry concludes with a review of the ways in which induction has been used in the domain of the social sciences.

One important aspect of inductive logic is that observations of the natural world need to be performed objectively and systematically. Observations that are made through our senses should not be guided by preconceived ideas or assumptions.

Generalizations are made based on these observations, and such generalizations can be strengthened by further observations. Therefore, a theory to explain a phenomenon is formulated on the basis of generalization made using specific observations. One common example used to explain induction is the observation of ravens. If every raven that has been observed so far is black, we will be led to draw a conclusion that all ravens are black.

Scholars and philosophers have described in various ways the principles or steps in induction. Carl Hempel and Anthony O'Hear have identified four main stages in induction. The first step involves the observation and recording of facts. This is followed by careful analysis of these facts without being guided by any hypothesis or preconceived ideas. Generalizations are then made on the basis of the analysis of the observations, and finally, these generalizations are tested further.

Rom Harré has explained that the three principles of induction are accumulation, induction, and instance confirmation. The principle of accumulation refers to the view that knowledge grows through accumulation of attested facts, which does not alter previously known facts. The principle of induction emphasizes that inference of universal laws can be done on the basis of accumulated facts. Statements from observations and experiments can be used to infer a general law. The third principle explains that our belief in a law depends on the number of positive instances that have been observed. For example, the more black ravens we observe, the more willing we will be to accept the finding that all ravens are black.

Induction, particularly Baconian induction—a method developed by Francis Bacon (1561–1626), one of the founders of modern science—represents a popular conception of science and what scientists do; namely, rigorously making observations and conducting experiments. Through meticulous analysis of the data that have been collected, they produce new discoveries. Baconian induction starts with the process of making observations that are free from three idols, so called. The first of these, *idols of the tribe*, refers to the inclination to see more regularities in nature than there actually are. The second, *idols of the cave*, refers to personal weaknesses based on our likes and dislikes. The third, *idols of the marketplace*, refers to

the confusions that arise from our language and terminologies.

In Baconian induction, the idea is to gather information regarding a phenomenon that allows us to move from particular observations to a general conclusion. According to Bacon, generalizations from particulars to universals need to be made in an orderly manner. The Baconian method of induction starts with observations. These observations are then generalized to a lower level axiom. Gradually from these lower level axioms, one can move up to a universal generalization, which he considered the goal of scientific inference. This, together with focusing on positive and negative instances, will help to avoid hasty universal generalizations. He also argued that experiments are important in scientific discoveries because they allow us to control the conditions of observations. Observing a phenomenon without any control would limit the depth of one's observation. By conducting experiments, researchers are able to manipulate conditions and observe how these conditions affect the phenomenon being studied. This would not be possible without performing experiments.

Baconian induction presents an ideal scientific method that is designed to be free from bias and fully based on objective observations. This was extremely important, especially at a time when many theories were influenced by philosophy or religious beliefs. It was a way of ensuring that theories were empirically sound. In reality, this is not necessarily how many scientific discoveries have been made. Even though many theories have been built inductively, many were not or could not have been discovered in this manner. In addition to this, as some scholars have noted, intuition may play an important role in scientific discoveries.

Origins and Development

Aristotle (385–332 B.C.E.) is considered the first thinker to describe inductive logic. His form of inductive reasoning was called *enumerative induction*, whereby knowledge is produced from a limited number of observations.

Bacon, as noted above, was the first major philosopher who explained the method of induction, which is still regarded by many as the scientific method. The Baconian method was stated in his

Novum Organum, which was published in 1620. Bacon suggested a more complex and gradual process of moving from particulars to generalizations and rejected the Aristotelian methodology of syllogistic demonstration. Bacon argued against the process of simply accumulating knowledge by generalizing from observations. He felt that such generalizations do not have applicability beyond the observed instances. His is an eliminative method of induction and emphasized the importance of focusing on negative instances. Negative instances are important to avoid a hasty generalization based only on positive instances. In making observations, it is important to identify elements that are present in positive cases, absent in negative cases, and vary accordingly in other cases. After making such observations, an accurate conclusion can be drawn. Bacon argued that focusing on both positive and negative instances stops one from jumping to conclusions from a limited number of observations.

Bacon demonstrated this type of scientific procedure through his investigation of the nature of heat. The investigation of heat can be done by collecting instances of heat and collecting information regarding the characteristics of each type of heat (e.g., rays of the sun, fires). These are seen as the positive instances. The next step is to gather instances that are in some manner similar to these instances but are not considered as heat (e.g., rays of the moon, starlight). These are called the negative instances. We also need to look at other instances with various degrees of heat. After looking at all the characteristics that are present in the positive instances, absent in the negative instances, and vary with the varying degrees of heat, a conclusion can be reached regarding the nature of heat.

Two centuries later, John Stuart Mill further elaborated the process of induction by stating what is called *Mill's canons of induction*. The canons are elementary experimental methods to identify general laws by identifying causes and effects of a phenomenon. The fundamental principle of induction is the proposition that the course of nature is uniform and the purpose of science is to discover causal laws. There are five canons of inductions, namely, the *method of agreement*, the *method of difference*, the *indirect method of difference* or *joint method of agreement and difference*, the

method of residues, and the *method of concomitant variations*.

Based on the method of agreement, in an investigation if at least two instances of the phenomenon have only one circumstance in common, then that circumstance is the cause (or effect) of the given phenomenon. In the case of an instance when the phenomenon under investigation happens and in another instance when it does not occur and all circumstances are the same except one, the missing one is the cause or the effect of the phenomenon. This is called the method of difference. The third canon combines the method of agreement and difference. When there are a few instances when the phenomenon occurs and it has only one circumstance in common and a few instances when it does not occur and the common circumstance is not present, that circumstance is understood to be the cause or the effect of the phenomenon. The fourth canon, the method of residue, is based on previously established cause and effect. If we have established from previous induction the cause of some aspect of the phenomenon, then what is left will be the cause of the remainder of the phenomenon. The final canon, the method of concomitant variations, emphasizes the connection between two phenomena. When one phenomenon varies together with another phenomenon, one will be the cause of the other.

Critique of Induction

The Scottish philosopher David Hume (1711–1776) is one of the first inductive sceptics. The problems he highlighted on induction were profound and received much attention from many philosophers even though he did not use the term induction in his argument. His arguments were mainly against the use of induction to establish causal relations. He highlighted that the major problem with induction is the difficulty in justifying the process of induction via either a deductive or an inductive argument. Even though we often use our experiences to make inferences in everyday life, Hume believed that induction cannot be justified rationally. Using past successes of induction to justify the use of induction puts us in a circular argument. Inductive inferences cannot be justified logically because there is no clear logic to show how one moves from observations to

generalizations. Therefore, one cannot claim that theories are not directly derived from data because these theories are explanations invented to account for the collected data.

Subsequently, critics have argued that there is no logical way of establishing the validity of the generalizations made through induction because induction involves the process of making generalized statements from a finite number of observations. In induction, we are using finite observations to generalize on infinite unobserved objects. Induction involves a process of making a claim beyond the mere observations. The conclusions reached on the basis of observations may be true for the observations that have been made, but there is no guarantee they will be applicable to future observations. Generalizations concerning unobserved objects, or the future, can only hold true if we depend on the principle of uniformity of nature, and this principle has not been proven empirically. Our senses are able to confirm what has been observed in the past, but there is no way of knowing how the same phenomena will look in the future. Therefore, we are making generalizations based on an assumption that has never been proven. Going back to the example of the ravens, even if all the ravens that have been observed so far have been black, there is no assurance that the next raven to be observed will be black. There is no way of guaranteeing this unless we are able to make all possible observations related to a phenomenon, and this will never be possible.

The second major criticism relates to the number of observations used in induction. Because it is impossible to make all possible observations, there is a need to decide on the number of observations that are necessary to make generalizations. Even if we decide only to make observations that are relevant to our study, there is still a need to decide on which observations are relevant to the study. In order to make the investigation manageable and practical, researchers are forced to decide on what to observe and how much to observe. This brings us to the next problem of induction, which relates to objective observations.

Objective observations are observations that are made without any preconceived ideas as to what the relevant observations are. It is believed that preconceived ideas about observations will introduce bias and jeopardize the objectivity of

scientific enquiry. Carl G. Hempel argued that such a narrow conception of induction will not allow scientific investigations to be carried out. Even if such an investigation is started, it would be almost impossible to collect all facts without any preconceived idea regarding the facts or observations. If, for example, one would like to study sand, is it possible to observe every grain of sand or every characteristic of sand? Karl Popper argued that it is impossible to start a scientific investigation without some sort of theory or idea regarding the investigation, which will guide the observations made by scientists. Even if objective observations are possible, many inferences can be made from the same observations and all can be equally correct. So how does one decide which generalization is accepted while which is rejected? Related to this, Nelson Goodman adds another problem with induction that he refers to as the problem of *projectability*. He emphasizes the difficulty in generalizing from past observations because there is no guarantee that the future will resemble the past. In addition to this, there is no clear way of distinguishing what is considered a confirmable generalization with a nonconfirmable or accidental generalization. He explained that although seeing another black raven can help confirm our generalization that all ravens are black, seeing a man who is a third son in a room does not confirm that all men in that room are third sons or that a man we would see in that room the following day will be a third son. The former is an example of a lawlike statement, while the latter is an example of accidental generality. Goodman differentiated that an accidental generalization often involves spatial and temporal restrictions (e.g., people in a room), whereas a lawlike generalization concerns complete generality (e.g., all ravens). Even though complete generality is a sufficient condition for lawlikeness, he also highlighted that it is often not easy to define complete generality. This he called the *new riddle of induction*.

On another front, Rom Harré argued that the growth of science happens not merely by the accumulation of facts or data by scientists as suggested by inductivists. The growth of science happens through the accumulation of facts and advancement in theory. Therefore, the explanation of a phenomenon cannot be achieved merely by

establishing patterns based on a finite number of observations. These observations may help to explain patterns of association but may not be able to provide an explanation on how these patterns occur. Furthermore, scientists have emphasized that many discoveries are made with the help of intuition. There are also some discoveries that occurred accidentally.

Justification of Induction

Since Hume's critique, many philosophers have tried to justify induction. Among them are Hans Reichenbach, Frederick L. Will, Peter Strawson, Bertrand Russell, and Jin Yuelin. They used various ways of justifying induction.

Reichenbach offered a pragmatic justification for induction. He accepts that induction cannot be justified inductively and deductively but questions whether it is necessary to have a justification of induction in order to accept that its conclusions are true. For him, while our ability to prove the truth of a conclusion will justify induction, such justification does not necessarily prove that the conclusion is true. Therefore, the proof that the conclusion is true is only a sufficient condition for the justification of induction. He believed that induction is the better way of reasoning based on experience compared with any other method of finding the truth. He uses the analogy of a cancer patient to emphasize his point: The patient will die unless the patient undergoes an operation. While there is no guarantee that the operation will be successful, it is the patient's only hope of survival. Therefore, the patient will choose the surgery because that is the best bet to survive. Similarly, Reichenbach believed that induction is our best bet and if there are laws of nature and we follow them, we can, to some extent, predict the future. Reichenbach's view has been criticized on the grounds that it is not an epistemic justification and does not give a reason to believe that induction is indeed true.

Peter Strawson gives a dissolution argument to justify induction. He argues that the demand for justification of induction is improper and a misunderstanding of induction. He rejects the idea of deductively justifying induction. According to him, if inductive argument is proved to be deductively valid, then it would be a deductive argument. This he finds absurd and stresses that induction must be

assessed using inductive standards. He disagrees that induction is justified by its success but says that a successful recipe for drawing conclusions regarding the unobserved must be one that has inductive support. The more success we have by repeating a recipe is the evidence we have in its favor.

Bertrand Russell maintains a contradictory position in relation to induction. While he accepts Hume's problem with induction, he also acknowledges that the importance of habit, knowledge, and common sense based on induction are rather reliable. He agrees that certainty may never be met through induction. He explains in his principle of induction that probability is the main aim of induction. If A is found to be associated with B, the more cases that show the association between A and B, the higher the probability A is always associated with B. After a certain number of observations, this probability moves closer to certainty. He argues that probability is always dependent on data. Additional data can alter the probability of an event. He uses the example of white swans to explain this point. If all swans that have been observed so far are white, it is probable that all swans are white. When a black swan is observed, it alters the probability of the proposition, but that does not mean that the original estimation was wrong because probability is always based on the available evidence. Later, he suggested five postulates that are necessary and sufficient conditions for validating scientific inferences. Jin Yuelin, a Chinese logician and philosopher who was influenced by Russell, shares similar views to Russell's but asserts stronger criticisms toward Hume's ontological and epistemological assumptions.

Frederick Will questions Hume's assertion that there is no evidence that the future will be like the past; hence, one cannot make inferences about the future. As a reply to this, Will discusses the meaning of the future. The future does not always remain as the future, but, as time passes it can become the present or the past. When this happens, we have the opportunity to learn more about the laws that have been stated previously. He explains this by saying that in 1936, 1937 and 1946 is the future, but in 1946, 1937 to 1946 ceases to be the future. However, this argument by Will has been criticized for not addressing the real problem of induction.

Induction in the Social Sciences

Émile Durkheim (1858–1917) pioneered the use of induction in the social sciences, particularly in the field of sociology. He emphasized the importance of explaining social phenomena on the basis of empirical evidence collected through objective observations, unlike other social theorists of his time who explained social phenomena through philosophical and metaphysical theories. He emphasized the importance of producing theories based on observations. In his view, a social phenomenon has to be observed in a similar manner as natural phenomena. Social phenomena can be explained in terms of social forces, which work on individuals just as do the forces that work on natural phenomena. He believed that preconceived ideas may distort the observation of a social phenomenon; therefore, he emphasized the importance of setting aside preconceived ideas. He, however, believed in defining the social phenomenon under investigation in advance to ensure the collection of the relevant facts to explain the social phenomenon scientifically. According to Durkheim, observations that are free of preconceived ideas on a clearly defined social phenomenon will enable true generalizations to be made regarding the phenomenon.

A *pure* form of induction is impossible in the social sciences because most researchers are guided by some theoretical ideas. It is also impossible to observe a social phenomenon objectively without some idea of which concepts are being observed or how each concept is defined. The choice of concepts and how these concepts are defined are necessarily based on some preconceived ideas or theoretical background. Owing to these problems, social scientists sometimes adopt a revised inductive methodology.

Most social studies that use inductive logic focus mostly on describing a social phenomenon in detail. Even though such detailed descriptions may be highly useful in explaining a phenomenon, induction is not often used directly to explain social phenomena. Even if generalizations are made through induction, these generalizations are always tentative and may be subject to revision because further research may reveal a different result.

In addition to this, social scientists often adopt different ontological and epistemological

assumptions than those used in the classical version of induction. The most important difference in the adaptation of induction in the social sciences is that social science investigation involves some preconceived ideas regarding how a social phenomenon is defined and what needs to be observed. These definitions are usually clearly stated to allow others to evaluate the conclusions and replicate the findings. Using a qualitative approach in the social sciences sometimes reduces the preconceptions and theoretical baggage used in social research, but such studies often use very different ontological assumptions than induction.

Premalatha Karupiah

See also Deduction; Epistemology; Hypothetico-Deductivism; Induction

Further Readings

- Bo, C. (2012). Justification of induction: Russell and Jin Yuelin. A comparative study. *History and Philosophy of Logic*, 33(4), 353–378. <https://doi.org/10.1080/01445340.2012.680703>
- Blaikie, N. (2007). *Approaches to social enquiry* (2nd ed.). Polity Press.
- Blaikie, N. (2010). *Designing social research* (2nd ed.). Polity Press.
- Ducheyne, S. (2005). Bacon's idea and Newton's practice of induction. *Philosophica*, 76, 115–128.
- Goodman, N. (1999). The new riddle of induction. In L. P. Pojman (Ed.), *The theory of knowledge: Classical and contemporary readings* (pp. 459–462, 2nd ed.). Wadsworth Cengage.
- Harré, R. (1972). *The philosophies of science. An introductory survey*. Oxford University Press.
- Hempel, C. G. (1966). *Philosophy of natural science*. Englewood Cliffs.
- Ladyman, J. (2002). *Understanding philosophy of science*. Routledge.
- Medawar, P. B. (1969). *Induction and intuition in scientific thought*. American Philosophical Society.
- Milton, J. R. (1987). Induction before Hume. *British Journal for the Philosophy of Science*, 38(1), 49–74. <https://doi.org/10.1093/bjps/38.1.49>
- O'Hear, A. (1989). *An introduction to the philosophy of science*. Clarendon Press.
- Russell, B. (1999). On induction. In L. P. Pojman (Ed.), *Theory of knowledge: Classical and contemporary readings* (pp. 438–441, 2nd ed.). Wadsworth.

- Strawson, P. (1999). Dissolving the problem of induction. In L. P. Pojman (Ed.), *Theory of knowledge: Classical and contemporary readings* (pp. 455–459, 2nd ed.). Wadsworth.
- Will, F. L. (1999). Will the future be like the past? In L. P. Pojman (Ed.), *Theory of knowledge: Classical and contemporary readings* (pp. 442–449, 2nd ed.). Wadsworth.

LOGIC AND LANGUAGE

The word *logic* originated from the ancient Greek word *logos*, which means “the word” or “what is spoken.” The exact meaning of the term *logic* is debatable among philosophers, but it is often understood as the study of arguments. As a discipline, it was studied in several ancient civilizations of the East and the West. The Greek philosopher Aristotle (384–322 B.C.E.) established it as a formal discipline in the West, while logic was developed independently in ancient India by Jains and Buddhists without any influence from Greece. Development of Indian logic dates back to the 6th century B.C.E. Indian logic is grounded in the Vedic theory of knowledge. Its initial concentration was on the freedom of soul from matter through systematic reasoning. Indian theories of inference cannot be easily distinguished into the deductive or inductive categories. The study of Indian logic is based on the Anviksiki school of logic by Medhatithi Gautama, as well as the Sanskrit grammar rules by the ancient Indian scholar Panini, which also influenced the famous Sanskrit epic the Mahabharata, among other works.

Aristotelian logic by far has found the widest application in philosophy and mathematics. Aristotle’s work is contained in six texts collectively named the *Organon* (“tool” in Greek), reflecting his approach to logic. Aristotle’s major work on logic includes *Categories*, *On Interpretation*, *Prior Analytics*, and *Posterior Analytics*, in which he discusses his system for reasoning and for developing sound arguments. His tradition was taken forward and developed by Islamic and medieval European logicians in the mid-14th century. The famous Arabic authors include al-Farabi, Avicenna, and Averroes.

Subsequently, the study of logic became less popular until it was revived in the mid-19th century.

This period witnessed the rapid development of mathematical, or symbolic, logic. Toward the end of the 19th century, three overlapping traditions in logic could be distinguished. The first, the *algebraic school*, focused on the relation between correct reasoning and mathematical operations. The *logician school* codified the underlying logic of scientific discourse into a single system. The *mathematical school* is the amalgam of the branches of mathematics such as geometry, arithmetic, and set theory. Mathematical logic has now come to be a central tool of contemporary analytic philosophy.

Aristotle’s Logic or Traditional Logic

Since Aristotle is widely regarded as the founder of logic, with profound implications for modern-day theorizing, aspects of his work relevant for the present discussion will be described here. Aristotle was the first to analyze *syllogisms*, which remained an important component of logic until the second half of the 19th century. A typical syllogism comprises an argument with three parts, namely a major premise, a minor premise, and a conclusion. The first two premises share exactly one term and entail the conclusion. This common term could appear as a subject in one and predicate in the other premise (*first figure*), subject in both premises (*third figure*) or predicate in both the premises (*second figure*). The conclusion contains the two nonshared terms of the premises. A particular combination of premises and conclusion is called a *mood*. Syllogistic logic is closely related to deductive reasoning as described below. In its original form, quantified conclusive statements were inferred from two quantified statements or premises. Basically the syllogistic language is confined to four sentence types:

Universal affirmative (A): All A are B

Example: All birds have wings

Particular affirmative (I): Some A are B

Example: Some birds are reptiles

Universal negative (E): No A are B

Example: No birds are animals

Particular negative (O): Some A are not B

Example: Some birds are not herbivores

Table I The Square of Opposition at a Glance

	<i>Affirmative</i>	<i>Negative</i>
Universal	A	E
Particular	I	O

This leaves conventional labels like A, I, E, O. These code letters are derived from the vowels of the two Latin words *affirmo* and *nego*. From here, A and O propositions and E and I propositions are contradictory. Propositions are contradictory when the truth of one implies the falsity of the other, and vice-versa. For example, where *All cats are mammals* is true, then *Some cats are not mammals* must be false. Propositions A and E are contrary where both of them together cannot be true. Propositions I and O are subcontrary where it is not possible for both to be false. If the statement *All days are long* is false, then *Some days are not long* must be true. Two propositions are said to be in *subalternation* when the truth of the first, the *superaltern*, implies the truth of the second, the *subaltern*, but not conversely. A proposition stand in the subalternation relation with the corresponding I propositions. The truth of the A proposition *All cars are four-wheelers* implies the truth of the proposition *Some cars are four-wheelers*. In traditional logic, the truth of an A proposition implies the truth of the corresponding I proposition and the truth of an E proposition implies the truth of the corresponding O proposition. Consequently, the falsity of an I or O proposition implies the falsity of the corresponding A or E proposition, respectively.

There is also another way to distinguish these propositions, and that is by arranging them in the *square of opposition* (Table I). The square is a diagrammatic representation of the previously described relations between these four propositions.

Some modifications have been proposed to the original square of opposition and other changes to Aristotelian syllogism, including an inductive approach that involves experimentation. A detailed description of these, however, is beyond the scope of this entry.

Logic in Contemporary Times

In the current era, logic is understood as the science of reasoning or reliable inference—whereby a

mathematician is interested in mathematical concepts and relations, or a philosopher is concerned with correct reasoning. Logical arguments rely on deductive or inductive reasoning. Under deductive reasoning, or top-down logic, valid conclusions are reached on the basis of given statements or premises, whereas in inductive reasoning, conclusions are reached by generalizing from an initial set of information or data base. For instance, if one determines or calculates that the probability of developing a flu increases if one lives in a cold environment, then it is reasonable to conclude that flu may be caused by cold. Thus, formulating explainable theories is important for induction. On the other hand, deductive reasoning takes place with the help of natural language; hence, a relation between natural and formal language assumes importance. Two examples of deductive reasoning:

Premise 1: If dolls are made out of chocolate,
then baby Anna is happy

Premise 2: Dolls are made out of chocolate

Conclusion: Baby Anna is happy

Premise 1: If the younger sister is sleeping, the
older sister is playing

Premise 2: The younger sister is sleeping

Conclusion: The older sister is playing

The premises given in both the examples are not facts about the real world, but the conclusions drawn are true as the premises are true. This is an essential property of reasoning in logic where the validity of conclusion is dependent upon the premises. Contrast this with the conclusion drawn below which is not valid in nature:

Premise 1: It is sunny or it is not raining

Premise 2: It is sunny

Conclusion: It is raining

The present entry lays stress on natural language and deals with a series of theories and supporting research studies that support the idea that human

reasoning and logic are well connected wherein meanings of linguistic expressions correspond to expressions of classical logic. The focus is on symbols of classical logic such as $\vee$, $\wedge$, $\neg$, $\forall$ that correspond to English words *or*, *and*, *not*, *all*, and on the complex interactions of $\&$, *or*, *not*, and quantificational expressions like *every*. These are called *logical connectives* or *logical operators* and work to join sentences in either natural or formal languages. The thrust is on how humans are biologically endowed with the capacity to reason logically, a capacity that could be studied through human languages consequently lessening the divide between the study of classical logic and philosophy of mind.

Alignment between this specific aspect of language and logic could be understood in terms of historical developments. The Chomskyan revolution in linguistics occurring during the 1980s, based on the work of Noam Chomsky, changed the equation between logic and natural language, helping to bridge the gap between natural and formal language. Structural linguists aimed at taxonomic study of languages before the revolution took hold. The mathematical logic of the structural era maintained a mind- and speaker-independent notion of language that did not advance the cognitive and psychological orientation toward language. Newer orientations with respect to language painted a more logical picture of human reasoning.

Language and Logic

Philosophy of language and logic share an important intimate relationship, as stated previously. Language is used to express thoughts and logic deals with concepts or objects of thought. Thus the study of language and logic has assumed great importance in modern-day study of mind.

The literature on cognitive science and linguistics is rife with debates on whether formal logic as studied by mathematicians plays a role in human reasoning or thinking. Classic experiments conducted in the structural era during the 1960s and 1970s presented adult participants with an array of cards. Participants were told beforehand that each card contained a letter on one side and a number on the other side. Furthermore, the cards which had a vowel on one side definitely contained an even number on the other side. The task

of the participants was to find out the cards that violated the given rules. Deductive logic dictates that the right selection for such a violation would be cards with a vowel and cards containing the odd number. The correct answer was given by only a handful of the participants. When these abstract rules and problems were replaced by a concrete problem and rule that gave contextual information in another experiment, the performance of the participants improved. The adult participants could perform better because they could make use of the contextual information. Conclusions consistent with beliefs or world knowledge are easier to draw than conclusions that rely on abstract content. Thus, many researchers believe that humans do not reason in a logical manner, paving a way for disconnect between logic and human reasoning.

Logical nativism, a result of the Chomskyan revolution, was offered as an alternative to the nonlogical picture of human reasoning. Logical nativism explains that humans are endowed with the innate faculty for logic. This view has gained momentum by research on human languages, favoring close alignment between formal and philosophical logic. Work carried out under this tradition shows that linguistic expressions like *and*, *or*, *every*, and *some* carry the same meaning as they carry in formal logic. Therefore, these linguistic expressions are called *logical expressions* and their semantic knowledge or truth conditions form part of *universal grammar* (UG) that is accessible to all irrespective of the individual language used or spoken.

Consider the truth conditions for logical disjunction ($\vee$) corresponding to the linguistic expression *or* that produces a value of true if at least one of its operands or propositions (A, B) is true or when both are true. The English word *or* receives the same interpretation as is true of disjunction in classical logic and hence is labeled inclusive *or*.

Contrast inclusive *or* with *exclusive-or*, which is false when both propositions (A, B) are true. The conjunction operator $\&$ corresponding to the linguistic expression *and* has a different set of truth conditions associated with it, according to which a statement is true only if both of the operands (A, B) are true. This is also called *logical conjunction* ($\wedge$). Table 2 details the values for conjunction:

Table 2 Truth Conditions for Logical Conjunction
($A \wedge B$)

Operands or Propositions		
A	B	Interpretation
True	True	True
True	False	False
False	True	False
False	False	False

Now, assuming that disjunction is inclusive-or in English, ($A \vee B$) is false when A and B are both false. However, the introduction of negation reverses truth-values, yielding one of de Morgan's laws of propositional logic where negated disjunctions generate a *conjunctive* entailment. Thus, $\neg (A \vee B)$, where $\neg$ is the symbol for negation (Table 4), could be rewritten as $(\neg A \ \& \ \neg B)$ or $(\neg A \wedge \neg B)$. This is how negated disjunctions work in the English language. Take a simple example in English like, *John did not speak Spanish or French* that could be understood as *John did not speak both the languages*.

A classic study looking at this phenomenon in English-speaking children had test sentences like *The girl who stayed up late will not get a dime or a jewel* versus *The girl who didn't go to sleep will get a dime or a jewel*. Only the former sentence represents $\neg (A \vee B)$; a look at Table 3 will help in understanding the associated truth conditions:

The results show that young children were able to grasp the idea in the statement representing $\neg (A \vee B)$ that was only evaluated as true when the girl did not receive a jewel and the girl did not receive a dime (highlighted in Table 3).

Evidence From Cross-Cultural Child and Adult Language Research

Arguments favoring disconnections between logic and human reasoning would show that sometimes adults would interpret negated disjunctions in a way that do not appeal to classical logic. For instance, speakers of languages like Japanese or Chinese would understand a statement containing negated disjunctions in a differential manner. For the example sentence, *John did not speak Spanish or French*, these speakers are more likely

Table 3 Truth Conditions Associated With $\neg (A \vee B)$

Operands or Propositions		
A	B	Interpretation
True	True	False
True	False	False
False	True	False
False	False	True

Table 4 Truth Conditions Associated With Logical $\neg$

A	$\neg A$
True	False
False	True

to interpret that *John either did not speak Spanish or he did not speak French*. This interpretation arises as the structural arrangement of negation and disjunction in Japanese and Chinese sentences, and is not comparable to English sentences. In Japanese and Chinese, disjunction is interpreted as taking scope over negation or the wide scope disjunction interpretation. This could be rewritten as $\neg A \vee \neg B$.

Even for adult speakers of English, ($A \vee B$) is subject to pragmatic or experience-based inferences implying that the disjunction of these two operands is understood as *not both*. As a result, adults would avoid using a statement like *John speaks Spanish or French* as a description of situation where *John speaks both the languages*. This is where adult reasoning does not conform to classical logic. Unlike adults, child speakers of English understand the English *or* as consistent with classical logic.

In order to show that disconnections between logic and human reasoning are only the result of experience or input, it is important to show that children irrespective of their native language community understand *or* in terms of classical logic. This is because all children have access to the linguistic principles of UG. Therefore, to begin with, young children across the world interpret *or* in a similar manner despite idiosyncratic language differences that may be absorbed later in the course of language acquisition consistent with the input from the environment. Children are expected to reason more logically than adults precisely owing to innate constraints in place.

Confirming evidence accrues from a developmental study conducted on 5-year-old Japanese-speaking children versus adults. In critical trials, both groups of participants judged whether a Japanese test sentence containing negated disjunctions would be true if one of the individual hypothetical situations corresponding to two operands (A, B) was true and the other was false. The judgments of Japanese adults and children differed where adults accepted the test sentence and children rejected it. Japanese children thus behaved like English-speaking children and adults. The proposed *semantic subset principle* explains this finding. The subset principle forces children to adopt the subset value of a parameter as their initial setting, and *inclusive-or* is one of them. The value is abandoned later only on the basis of positive evidence available in the local language. This makes sense because if children adopted the superset value, they would generate sentence meanings that are not licensed in the local language, in addition to ones that are licensed.

Details of Good Testing

To reiterate, logical expressions or words (e.g., *or*, *every*) are not only subject to inherent semantic knowledge, as has been shown in the case of children, but also subject to pragmatic influences. This gives rise to conversational implicatures wherein a hearer interprets that a speaker using a weaker term like *or* was not in a position to use a stronger term like *and*. An important aim for any testing exercise geared toward uncovering details of innate semantic principles is to build situations where these implicatures are canceled out. This section will focus on one such testing scenario. A context where these implicatures are canceled out is called the *prediction mode*. For instance, consider a sentence like *There will be a blue or red car in the showroom*. Adults would consider a statement like this to be true even if both types of car are found at the showroom after the sentence is uttered. This happens because some of the circumstances linked with the English word *and* are also linked with the English word *or* as noted previously.

The prediction mode introduces uncertainty about the outcome, a feature that could be fruitfully employed for making the testing conditions

felicitous. The felicitous condition is based on conversational implicatures where a speaker is justified in using the word *or* if he/she is unsure about what will happen in the future. This could be successfully employed for testing children for the purpose of uncovering innate semantic principles. Describing these testing procedures in detail deserves an important mention, as some of the work done could lead to the erroneous conclusion that children's reasoning does not conform to classical logic as has been shown before for adults, or in other words that for children English *or* corresponds to an *exclusive-or* reading. Such results are the fallout of experimental manipulations biased toward exclusivity. Consider situations where children were asked to pick either a red balloon or a blue balloon. Children under these testing conditions were actually found to pick the red or the blue balloon more often rather than choosing both colors together, implying that they cannot access conjunctive truth conditions associated with the English *or*, consistent with classical logic.

The previously described research thus suffers from two major drawbacks: first, that children were required to initiate a physical act which in itself has limited value when it comes to gauging the range or scope of truth conditions associated with the linguistic expression *or*; and second, these testing scenarios gave rise to conversational implicatures. The child may assume that if the adult tester wanted them to pick both types of balloon, she would have used *and* instead of *or*. Thus, it becomes critical to cancel these conversational implicatures in order to study the range of truth conditions associated with English *or* and accessed by children. Moreover, the very nature of the initiation of physical act shows that children only have access to one grammatical interpretation and these acts cannot be used to gauge whether children would show evidence for another interpretation in a different situation.

One such technique that takes into account all the good testing ingredients as mentioned earlier is called the *truth-value judgment task* (TVJT). In a typical version, two experimenters participate where one is responsible for presenting a short story with the help of toys and props and the other is responsible for manipulating a puppet that witnesses the whole story along with the child

who is being tested. At the end of the story, the puppet summarizes the events and presents a critical statement to be evaluated by the child. The child's task is to reward the puppet if the critical statement appears to be the best description of the story events and to punish the puppet if the critical statement is not an accurate description. The way the puppet is judged reveals whether the child has access to the range of truth conditions associated with English *or*. This can be achieved by manipulating experimental or story contexts, each associated with a particular truth condition. Children whose grammar lacks the inclusive-or interpretation should judge the presented critical statements as false in story contexts where both the operands (A, B) are true, whereas children whose grammar allows the inclusive-or reading should judge the same statement as true.

In a classical experiment conducted in the 1990s, children were shown to indicate knowledge of conjunctive truth-conditions associated with English *or* when conversational implicature of exclusivity was canceled out with the help of the prediction mode. English-speaking children were tested in the prediction and the description mode. In a description mode, the puppet simply described the events that had already taken place. Consider the previous example where a puppet might say, *There will be a blue or red car in the showroom* after, for example, the children had been shown a toy showroom containing both types of cars. Under such an example scenario, children in the prediction mode did much better than the children in the description mode.

Another testing scenario where the conversational implicature of exclusivity could be canceled is the use of the verb *can* in a sample statement like *Mary or Jane can finish the painting*. If this statement is understood as a generic one about the ability of either Mary or Jane to complete a painting, then the inclusive-or reading becomes available. Experimental manipulations like this in the past have also shown that children have access to the inclusive-or reading.

A glimpse of the following tables would make the differences clear where the crucial difference between inclusive-or (Table 5) and exclusive-or (Table 6) reading is highlighted in bold:

Table 5 Inclusive-or Reading as It Conforms to Classical Logic

<i>Operands or Propositions</i>		
<i>A</i>	<i>B</i>	<i>Interpretation</i>
True	True	True
True	False	True
False	True	True
False	False	False

Table 6 Exclusive-or Reading Also Arising From the Conversational Implicature of Exclusivity

<i>Operands or Propositions</i>		
<i>A</i>	<i>B</i>	<i>Interpretation</i>
True	True	False
True	False	True
False	True	True
False	False	False

An Explanatory Hypothesis: Modularity Matching Model

The Modularity Matching Model (MMM) described in this section serves as an organizing framework for explaining how and why a cognitive system (e.g., humans) would assign different reading to linguistic expressions when the expressions are subject to conversational implicatures. The MMM borrows theory from UG and within it the role ascribed to experience is minimal, at least in the acquisition of syntactic and semantic principles. The role of experience is limited only to setting parameter values as explained for the semantic subset principle. The model rests on two assumptions: first, that the operation of the language faculty is governed by specific principles which are not shared by other systems such as vision and memory. Second is the assumption of continuity wherein both the child and the adult language-processing systems are the same, or in other words the principles of the semantic or syntactic components are the same. According to the MMM, operations within the language faculty are organized hierarchically where the results of the analysis taking place at the lower levels are transferred upward for higher levels of analysis.

Operations taking place at the higher level components are called the *pruning principles*. These pruning principles apply to the output or structural descriptions of the lower level components. In cases where multiple outputs are transferred for higher levels of analysis, the pruning principles will maintain or reject these. Under such a scenario, pruning principles cannot influence the interpretations of already unambiguous sentences. For the current purposes, pragmatics or conversational implicatures can be understood as these pruning principles, while the analyses taking place at the semantic level would be lower in the hierarchy. In case of children who have yet to acquire the pragmatics, only one output from the semantic level dominates, which in this case would be the inclusive-or interpretation. The experimental manipulation described in TVJT was aimed at constraining the functioning of the higher level pragmatic components so as to understand the operations taking place at the lower semantic or syntactic levels.

Negation, Disjunction, and Quantificational Words

The study of logic and the study of formal semantics, or logical semantics, are closely tied. *Logical semantics* deals with the notion of entailments (also called *logical consequence*), which is a fundamental concept in logic. Entailment relations are relations between two statements and are considered true when one logically follows from the other within the framework of deductive reasoning. Consider the premises and conclusions explained previously, where conclusions are considered valid and logical if they follow from premises. Therefore, another exemplar from semantics deserves a mention. Moreover, this description will help to strengthen the arguments advanced in the previous section in favor of our innate faculties, with more focus on negation and disjunction as these interact with other phrases such as universally quantified noun phrases (NPs) (e.g., *every*) in the subject position of sentences.

To gauge how the interpretation of *every*, another logical expression, changes when it occurs in the object versus subject position, consider the following sentences: *Every researcher didn't write the paper* versus *The researchers didn't write every paper*. The interpretation of the first sentence is

ambiguous. On one glance, it could be understood to mean that every researcher did not write the paper or none of them did. This is the isomorphic reading, where the overt positions of negation and quantified NPs are aligned with their interpretation. Or in other words *every* is interpreted outside the scope of negation (*every* > *not*). On the other hand, it could be understood as some or at least one researcher writing the paper. This is the nonisomorphic reading, where *every* is interpreted inside the scope of negation (*not* > *every*). However, the interpretation of the second sentence is not ambiguous. Here, *every* is always interpreted within the scope of negation. This sentence could only mean that all the researchers did not write all the papers. Therefore, the interpretations of the interactions between negation and universally quantified NPs vary, depending upon whether *every* occupies the subject or the object position.

Other interactions have been noted between disjunction and the universal quantifier (Table 7). A sentence like *Every linguist who studies English or Hindi got the award* could be interpreted to mean that every linguist who studies English got the award and every linguist who studies Hindi got the award. Here, disjunction appears in the subject phrase of the universal quantifier. Similar interpretation for interactions between *every* and disjunction have been noted for languages like Chinese and Japanese. Compare this sentence with a sentence like *Every linguist who got the award studies English or Hindi* which could only be interpreted to mean either every linguist who got the award studies English or every linguist who got the award studies Hindi, but not both. Children from an early age are shown to differentiate between the sentences where the arrangement of disjunction and *every* gives rise to differing interpretations, thus yielding another linguistic principle where disjunction gives rise to a conjunctive interpretation when it occurs only in the

Table 7 Interactions of *Every* With Or

Statements	Interpretations
Every SUBJ [.....Or.....] PRED [.....]	Conjunctive
Every SUBJ [.....] PRED [..... Or.....]	Disjunctive

subject phrase of the universal quantifier. However, disjunction will give rise to a disjunctive interpretation when it occurs only in the object phrase of the universal quantifier. The following table underscores this difference:

It is important to point out that the conjunctive-versus-disjunctive interpretations arise as English *or* is inclusive in nature, corresponding to the interpretation it receives in classical logic. Thus, De Morgan's laws (after the 19th-century British mathematician Augustus De Morgan) are preserved not only at the level of logic but also for semantic interpretation in human languages.

Evidence From Cross-Cultural Child and Adult Language Research

Studies conducted on preschool children using TVJT show that their interpretation is not aligned with those of the adults in case of sentences when it comes to interactions of *every* in the subject position with that of negation. The way these sentences are structured can lead to ambiguities while interpreting quantificational expressions, because processing can occur at two different levels, namely syntax and semantics, as explained above for the MMM. Semantics is higher up in the hierarchy than syntax. Consider a phrase like *Everyone didn't read* could be interpreted as either *no-one read* $\forall x$ [person (x) $\rightarrow$ $\neg$ read (x)] or perhaps as that *some people did not read* $\neg \forall x$ [person (x) $\rightarrow$ read (x)]. In the former interpretation, based on syntactic level, the universal quantifier *every* takes scope over negation (*didn't*) while in the latter, negation takes scope over *every*. Therefore, processing taking place at these two levels may not be isomorphic in nature. Adults usually get the latter reading while children aged 4 years usually get the former.

Sentences like these were tested in a well-known experiment where the testing context involved three horses jumping over a fence. Two horses could successfully jump over the fence but the third horse could not. Toward the end, the puppet summarized and said *Every horse didn't jump over the fence*. The statement is true on the adult reading but false on the understanding that children usually show. But this asymmetric result gives rise to an important question on how

children eventually develop an adult level understanding. This question assumes importance in the light of a caveat where the semantic knowledge or truth conditions of the *none* reading satisfies or entails the truth conditions of the *not all*, or adult reading. Or in other words, if it is true that none of the horses jumped over the fence, then it also true that not all of them did, but not the other way round. Careful experimentation revealed that adults could be made sensitive to the *none* reading by suitable testing contexts.

Coming back to the same question of how children eventually converge on an adult-level understanding, they only need to add to their grammatical repertoire based on the linguistic input they receive from adults again implying compatibility with the innate principles of UG. Therefore, it makes sense to start with the dispreferred interpretation consistent with the already explained semantic subset principle. This line of reasoning is grounded on support from cross-linguistic literature such as the Chinese, where sentences like *Every horse didn't jump over the fence* are not ambiguous. Sentences like these only give rise to an isomorphic reading in Chinese. Adult Chinese speakers come up with the same interpretation as English-speaking children. Thus, assigning an English-speaking adult-like interpretation seems to be another case of parametric variation, where all children would seem to start with a subset value or an isomorphic reading corresponding to Chinese adult-like understanding. Starting with a superset value poses a learnability problem whereby children may never come across evidence or negative evidence that will be helpful in setting a native-like parameter value.

Atypical Population

In order to provide stronger support for the hypothesis of logical nativism, interpretations aligned with the syntactic and semantic levels in the MMM should accrue from developmental disorders (like autism) where linguistic knowledge has been argued to be developing atypically as compared to typically developing children. Autism is a part of a spectrum that is labeled as the autism spectrum disorder. It is a lifelong developmental disorder markedly characterized by difficulties with language and communication. A very recent

study was conducted on participants diagnosed with autism with a mean of 16 years and 3 months on *Every . . . not* sentences. The results show that the performance of these participants was very similar to a 4-year-old control group of typically developing children. The results hint at the possibility that even at an older age, participants with autism are unable to interpret these sentences in an adult-like manner, implying that the basic syntactic and semantic principles are intact. What seems to be disturbed is the acquisition of pruning principles, which comes later in development.

Final Remarks

A disconnect between logic and human reasoning seems to be a result of the way research was conducted on human logic in the past. Classical experiments reported in the 1960s and 1970s showed that human reasoning does not conform to the laws of logic. This has since been challenged with the help of work and developments taking place within the nativist approach to (or the Chomskian view of) language. With this approach, language acquisition is explained as a biologically driven process where learning occurs as a by-product of task-specific computational mechanisms. To reiterate, children are biologically endowed with core principles (common to all human language) of UG that are not dependent upon learning experience, as shown with the help of work on logical operators (*every*, *or*, *not*). The crux of UG dictates how human languages differ from each other. These variations are called *parameters*. As a result of this set-up, learning of language proceeds rapidly and effortlessly. The studies reported here provide evidence for a connection between logic and human thinking whereby human reasoning confirms to the laws of logic. Language is seen as an important window to this connection, favoring close alignment between formal and philosophical logic gauged through its study in children and adults with a focus on syntactic and semantic components.

Neha Khetrpal

See also Logic, Formal and Informal; Logic, Inductive; Logical Theory; Pragmatism

Author's Note: I am deeply indebted to my principal supervisor, Associate Professor Rosalind Thornton, for helping me embark on this new terrain of research in linguistics. I also wish to acknowledge the financial support provided to me by International Post Graduate Research Scholarships (IPRS).

Further Readings

- Chierchia, G., & McConnell-Ginet, S. (2000). *Meaning and grammar: An introduction to semantics*. MIT Press.
- Chierchia, G., Crain, S., Guasti, M. T., & Thornton, R. (1998). "Some" and "or": A study on the emergence of logical form. *Proceedings of the Boston University Conference on Language Development*. Cascadia Press.
- Chomsky, N. (1986). *Knowledge of language: Its nature, origin and use*. Praeger.
- Crain, S. (2012). *The emergence of meaning*. Cambridge University Press.
- Crain, S., & Thornton, R. (1998). *Investigations in universal grammar: A guide to experiments on the acquisition of syntax and semantics*. MIT Press.
- Gualmini, A., & Crain, S. (2002). Why no child or adult must learn De Morgan's Laws. *Proceedings of the 24th Annual Boston University Conference on Language Development*. Cascadia Press.
- Honderich, T. (2004). *The Oxford companion to philosophy*. Oxford University Press.
- Johnson-Laird, P. N. (2006). *How we reason*. Oxford University Press.
- Leszl, W. (2004). *Aristotle's logical work and his conception of logic*. Topoi.
- Musolino, J., Crain, S., & Thornton, R. (2000). Navigating negative quantificational space. *Linguistics*, 38(1). <https://doi.org/10.1515/ling.38.1.1>
- Noveck, I. A., Guelminger, R., Georgieff, N., & Labrueyere, N. (2007). What autism can reveal about every . . . not sentences. *Journal of Semantics*, 24(1), 73–90.
- Staal, F. (1988). *Universals: Studies in Indian logic and linguistics*. The University of Chicago Press.
- Tennant, N. (2014). Aristotle's syllogistic and core logic. *History and Philosophy of Logic*, 35(2). <https://doi.org/10.1080/01445340.2013.867144>
- Tomasello, M. (2000). First steps toward a usage-based theory of language acquisition. *Cognitive Linguistics*, 11(1–2), 61–82.
- van Benthem, J., & Ter Meulen, A. (2011). *Handbook of logic and language*. Elsevier.

LOGICAL THEORY

The word *logic* is generally used in two different ways. We use it to say such things as “Her logic seems faulty to me,” or “I disagree with his logic.” In these instances, we are using the word *logic* to refer to a person’s reasoning skills. However, the word *logic* can also refer to an academic subject that is defined as *the systematic study of the standards of good reasoning*. Humans have been thinking logically since prehistoric times, but the academic theory dates back to around the 5th century B.C.E. in Athens, where ancient Greek philosopher Aristotle taught logic classes and wrote treatises on logical theory. The new academic subject *logic* was not focused on reasoning about particular topics but was rather an examination of the standards any kind of reasoning must follow if it is to be considered *good reasoning*. This entry traces the origin and development of logic from ancient Greece through to the present day, including an account of arguments, the varieties of logic, the principles of truth and validity, and the contributions of major thinkers over the century.

At Aristotle’s time, the academic subject we now know as logic did not have this name, and it was only decades after his death that the term *logos* was used, attributed to a group of philosophers known as the *Stoics*.

The Birth of Logic

Throughout history, cultures have embraced intricate systems of reasoning with logical methods evident throughout all human thought. However, an explicit analysis of the principles of reasoning was only developed in three traditions, those of China, India, and Greece. Of these, only the logic descending from the Greek tradition, and in particular Aristotelian logic, has found wide application in mathematics and science. In ancient Greece, logic was known as *dialectic* or *analytic*.

Aristotle’s logic was developed further by Islamic and medieval European logicians and reached a high point in the middle of the 14th century. It was then several centuries later that logic became reinvented into an exact method of proof in mathematics. The development of modern-day *symbolic* or *mathematical* logic is one of

the most significant in its 2000-year history and arguably one of the most important events in intellectual history. Progress in mathematical logic in the first part of the 20th century, particularly in the works of Kurt Gödel and Alfred Tarski, had an impact on analytical philosophy and philosophical logic. Subjects such as modal, temporal, deontic, and relevance logic came into being.

Aristotle was born in Macedonia, to the east of the Greek peninsula. He was later sent to Athens to study at the academy of the famous philosopher Plato. The Greeks took the name philosophy from *philo*, to love, and *sophia*, wisdom: that is, the love of wisdom. The Greeks had already started questioning the myths and traditional explanations of the universe in written works and in recorded discussions, and they pioneered a radical new way to make sense of the world. They rejected myths and the pronouncements of authority on the basis that there was no *good reason* to accept their claims as true; that is, in correspondence with reality. In place of mythical stories and pronouncements, they looked for explanations that were based on unaided reason and on observation, both of which could be made by anyone. They set down their theories along with supporting evidence for their theories in writing and passed it around, so others could comment, discuss, and debate their findings. Theories were proposed and criticized, defended or revised, and rejected on the basis of reasoning and the available observable evidence—and without reference to myth or questionable statements of authority.

Thus, the first philosophers looked for rational explanations of the world and of the things within it. These accounts had to be justified based on reasoning and evidence alone. Ancient Greece was the birthplace of the philosophical tradition and created an intellectual revolution. Aristotle decided to make it his job to create a separate discipline that would study the universal standards of reason itself rather than reasoning about this or that particular theory. His theory of pure reason would help to sort out the complications for finding reasons for and against each and every theory. He went on to write many treatises on logical theory, known collectively as the *organon*, and much of it is as valid today as it was over 2,000 years ago. Academics still devote their lives to studying Aristotle’s works on logic, and new books are still appearing year after year.

The Basics of Logical Theory

Logical theory begins with the concept of an argument. An argument does not in this case involve any form of heated emotional dispute. When used in logic, it means *reasoning put into words*. For logicians, when you put reasoning into words, an argument is the result. When it comes to defining logical theory, a more detailed exposition is usually required. This generally involves the statement that an argument consists of one or more statements, termed premises, offered as a reason to believe that a further statement, termed the conclusion, is true—that is, it corresponds to reality. In logic, the word argument means *one or more premises offered as reasons or as evidence for the truth of a conclusion*.

When we are listening to an argument, it can be difficult to ascertain which statements are the premises and which are the conclusion. For this reason, the English language contains what logicians term *argument indicator words*. In order to let the audience understand when to expect the conclusion, the statement should be introduced using a word or phrase such as *in conclusion*, *therefore*, *thus*, *consequently*, and so on. *Premise indicator words* and *conclusion indicator words* enable a reader to follow the progression of reasoning.

Deductive and Inductive Arguments

Following in Aristotle's footsteps, logical arguments are divided into two main types: deductive and inductive. A *deductive argument* sets out to show that its conclusion must necessarily be true. That is to say, a deductive argument seeks to establish its conclusion conclusively. The implicit or explicit claim in a deductive argument is that if the premises are true, then the conclusion must absolutely be true too.

An *inductive argument* is an argument that seeks to show that its conclusion is probably true, although not necessarily. An inductive argument claims that only if the premises are true, the conclusion is likely or reasonably certain to be true.

Deductive arguments are made explicit by introducing the conclusion with words such as "therefore it must be true" or "it necessarily follows that" or "it is conclusively proven that" or

"therefore it is certain that" These types of phrases are known as *deductive indicators*.

An inductive argument is made explicit by introducing the conclusion with phrases such as "therefore it is likely that," "therefore it is probable that," or "therefore there it is reasonable to suppose that" These types of phrases are known as *inductive indicators*.

Valid and Invalid Deductive Arguments

When a deductive argument is successful in illustrating that its conclusion is true, if all its premises are true, it is termed a *valid* deductive argument. Deductive arguments that fail to show that the conclusions are true if their premises are true are termed *invalid* deductive arguments. Thus, a valid argument may be defined as a deductive argument, where it is the case that if the premises are true, then the conclusion must be true. An invalid argument may be defined as a deductive argument, whereby it is not the case that if the premises are true, then the conclusion must be true.

The following arguments are deductive since each aims to show that its conclusion must be true if the premises are true. Only the first argument is valid, however, with the second invalid. The arguments are deductive because each aims to show that all human beings are mammals:

1. All mammals have warm blood.
2. Therefore, it must be the case that all human beings have warm blood.
1. All human beings are mammals.
2. All cats are mammals.
3. Therefore, it must be the case that all human beings are cats.

This illustrates that not all reasoning is equal and that some reasoning is better than others. Validity is something that often puzzles us, since an argument can be valid even though it has false premises and a false conclusion.

Look at the following deductive argument:

1. All teenagers are millionaires.
2. All millionaires are Christians.
3. Therefore, all teenagers are Christians.

Even though the premises are false as well as the conclusion, the argument is a valid one. It is valid because if the premises were true, then the conclusion would have to be true too. The argument then fits the definition of what it is to be a valid argument. The premises are false, yet the argument is absolutely valid, which does not fit in with our common sense. This shows that in logic, true premises are not required for validity, and that the term *valid* does not equal true. An argument is valid just so long as it is the case that if the premises were true, then the conclusion would also be true. True premises are not necessary.

However, we naturally want more than validity from a deductive argument, we want the truth. When an argument is valid, and its premises are all true too, then the argument is termed a *sound* argument. A sound argument features two characteristics:

1. All of the premises are true.
2. It is a valid argument.

Since truth corresponds with reality, it is the ultimate goal of good reasoning. Thus, it follows that soundness, rather than simply validity, is the ultimate goal of a deductive argument. For reasoning to be worthwhile the deductive argument has to be sound as well as valid. Consider the following argument that is both valid and sound:

1. All whales are mammals.
2. All mammals are animals.
3. Therefore, it must be the case that all whales are animals.

The argument that was previously considered, involving teenagers and Christians, was a valid argument but it was unsound.

Strong, Weak, and Cogent Inductive Arguments

Inductive arguments that succeed in showing that their conclusions are probably but not certainly true, if the premises are true, are termed a *strong inductive argument*. Inductive arguments that fail to show that the conclusion is probably but not certainly true, if the premises are true, is termed a *weak inductive argument*.

So, it follows that a strong argument may be defined as an inductive argument, whereby it is

the case that if the premises are true, then the conclusion is probably true. A weak argument may be defined as an inductive argument, whereby it is not the case that if the premises are true, then the conclusion is probably true.

The following arguments are inductive, since each attempts to show that its conclusion is probably but not certainly true. The first is strong and the second is weak:

1. In all recorded history, it has never snowed more than 6 inches in Michigan in July.
2. Therefore, it will probably not snow more than 6 inches in Michigan in July.

1. Jan is a member of the Republican Party.
2. Some known communists have been members of the Republican Party.
3. Therefore, Jan is probably a communist.

This again goes to show that not all reasoning is equal and that some arguments are superior to others. An inductive argument may be strong even though it has false premises and a false conclusion. For example:

1. For the past 5 months, it has been snowing in Michigan every day. It is below 40° in Michigan, and the sky in Michigan is packed with snow clouds.
2. Therefore, it will probably snow today in Michigan.

While the premises are obviously false, as is the conclusion, the argument is a strong one. It is strong because if the premises were true, then the conclusion would probably be true as well: If the premise is true, then the conclusion is likely to be albeit not certainly true, so the argument fits with the definition of a strong argument.

As before, however, strength is not sufficient for an inductive argument; we also require truth. If an argument is strong and its premises are true, then the argument is termed a *cogent argument*. A cogent argument has two properties:

1. All the premises are true.
2. It is strong.

Truth is the ultimate goal of reasoning, and cogency is the ultimate goal of inductive arguments.

The following argument is simultaneously strong and cogent:

1. Most cars run on gasoline.
2. The president's limousine is a car.
3. Therefore, the president's limousine probably runs on gasoline.

The previous example that involved Michigan and snow was strong, but it was not cogent.

Consistency, Implication, and Equivalence

Using reason, we have judged deductive arguments as either valid or invalid and assessed the strength of inductive arguments. We also need to use reason to help us decide whether or not two of our beliefs are logically inconsistent or in conflict, and reason will also help in identifying logical relations between beliefs. It is because of this that logical theory also seeks to examine the logical relationships that exist between *declarative* statements and statements' logical properties. There are four terms of special importance: *consistency*, *inconsistency*, *implication*, and *equivalence*.

Consistency and Inconsistency

Two statements are consistent if (and only if) it is possible both are true. Two statements are inconsistent if (and only if) it is not possible both are true. For example, the following statements, given their standard meanings, are consistent:

1. Anna is 11 years old.
2. Anna is a violin player.

And the following statements, given their standard meanings, are inconsistent:

1. Anna is 11 years old.
2. Anna is a teenager.

Implication

One statement implies a second statement if (and only if) it is not possible that the first statement is true and the second is false. Put another way, a statement S implies a statement P when, and only when, it is the case that if S is true, then

P is true. For example, in the following case, the first sentence implies the second:

1. Elizabeth is 14 years old.
2. Elizabeth is older than 10.

However, in this next case, the first sentence does not imply the second:

1. Elizabeth is a netball player.
2. Elizabeth is a millionaire.

Equivalence

Two statements, S and P, are equivalent if (and only if) S implies P and P implies S. Put another way, the two statements are equivalent when and only when it is not possible that they differ in terms of truth and falsity: If one statement is true, then the other is true. If one statement is false, then the other is false. In the following example, the two statements are logically equivalent:

1. Elizabeth is taller than Anna.
2. Anna is shorter than Elizabeth.

These two statements on the other hand are not equivalent:

1. Elizabeth is older than Anna.
2. Anna is 11 years old.

Necessity and Contingency

Our reason comes into play too when we decide if a statement is necessary or contingent, and logic is involved in defining the relevant terminology to ensure our thinking is as clear as possible here. There are four important terms to consider: *a necessary truth*, *a necessary falsehood*, *a contingent truth*, and *a contingent falsehood*. These are defined as follows:

A statement is necessarily true if it is in fact true and cannot possibly be false. That is, that the statement is true in all possible instances, and that there are no circumstances whereby it could be false. For the purpose of logical theory, *possible circumstances* can be defined as a circumstances whose descriptions are nonself-contradictory. This is the most encompassing concept of possibility conceivable to humankind.

A statement is necessarily false if it is false and there is no possibility of it being true. That is, that the statement is false in all conceivable circumstances and there are no possible situations in which it would be true.

For example, the following statements are all necessarily true, given their standard meanings:

1. All squares have four sides.
2. The derivative of a constant is zero.
3. Number 4 is greater than number 3.

According to their standard meanings, the following statements are of necessity false:

1. All squares have five sides.
2. $1 + 1 = 3$.
3. Number 5 is less than number 3.

Contingency

A statement is contingently true if it is true but there are possible circumstances in which it would be false. That is, it is true but may have been false should circumstances have been sufficiently different.

A statement is contingently false if it is false but there are possible circumstances in which it would be true. That is, it is false but may have been true if circumstances had been sufficiently different.

Contingent truths could be illustrated thus:

1. The Mamas and the Papas performed at Woodstock in 1969.
2. It is sunny in Detroit on September 22, 2010.

Examples of contingently false statements could be:

1. Bob Dylan played Woodstock in 1969.
2. Richard Nixon became president in 1960.

The Five Branches of Logic

Now that we have defined the main concepts of logic, we can take a look at the five important branches of logical theory. These are *categorical logic*, *predicate logic*, *truth-functional logic*, *modal logic*, and *informal logic*.

Categorical Logic

Categorical logic is concerned with finding the universal standards governing categorical arguments. Categorical arguments look like this:

1. All dogs are mammals.
2. All mammals have lungs.
3. Therefore, all dogs have lungs.

Aristotle was the first to discover and treat systematically universal standards of logical reasoning. He began by deconstructing categorical reasoning into essential elements and then studied how to fit the parts together to make the different forms taken by categorical reasoning. By making precise definitions for the types of sentences that make up categorical arguments and defining a categorical argument (syllogism) precisely, he discovered that there are four basic forms a categorical sentence can take. From this, he deduced that there are 256 basic forms of categorical reasoning, 24 of which are unconditionally valid reasoning patterns.

Aristotle devised from this basis a set of general algorithms capable of determining precisely in every possible case whether a syllogism is valid or invalid. (An algorithm is a precise rule for completing a task that guarantees a definite outcome, after the accomplishment of a finite series of steps.) Aristotle also developed a precise axiom system, which allowed him to prove, in a mathematically precise way, the validity or invalidity or every possible form of categorical reasoning.

In brief, a categorical argument is made up of categorical statements. Categorical statements express the relationships between two categories (groups of things) by stating that either all, none, or some of one category belong, or do not belong, to a second category. A sentence making up a categorical argument can have four forms. Supposing A and B stand for categories (of things), then we can say:

1. All A are B.
2. No A is B.
3. Some A are B.
4. Some A are not B.

This can be expressed as:

1. All dogs are mammals.
2. No dogs are reptiles.
3. Some dogs are pets.
4. Some dogs are not pets.

Every categorical sentence can be divided into four parts: quantifier, subject term, predicate term, and a copula. The subject term identifies the subject category or what the sentence is about, the predicate term is the identifier of the predicate category, while the quantifier states how many subject category members (of dogs) belong to or do not belong to the predicate category (of mammals). The copula (*are*) links subject and predicate and indicates whether the quantity refers to things belonging to the predicate category or those that do not belong to the predicate category.

Truth-Functional Logic

In truth-functional logic, arguments look like this:

1. If it is daytime, then the sun is out.
2. It is daytime.
3. Therefore, the sun is out.

This type of argument dates back to the Stoics of the 3rd century B.C.E. The logic was updated by the philosopher and mathematician Gottlob Frege in the latter part of the 19th century. Truth-functional logic concerns itself with declarative sentences that are compound if they contain one or more sentences within them or simple if they don't. A sentence operator is a word or phrase that links one or more sentences into a compound sentence. For example:

Plato taught at the Academy, and Aristotle taught at the Lyceum.

In this compound sentence, *and* is the operator joining the two simpler sentences together. A sentence has the truth-value of true if it expresses a true proposition, and a truth-value of false if it expresses a false proposition.

Predicate Logic

Predicate logic was invented by Gottlob Frege, who tried to construct an axiomatic proof originating with statements of pure logical theory. After following a set of deductively valid inferences, the proof would lead to the principles of pure arithmetic or number theory, which would eventually reach out to all mathematical principles. Frege's theory sought to bring new domains under the mantle of logic through generalization, a form of logic that is also called *modern-day* logic or *symbolic* logic." It is called this because it is a logic expressed in symbols and formulae that resemble those of mathematics, taking it beyond the scope of traditional logic. Frege invented the first *formal* language for logic theory, giving it a complete system of syntaxes and semantics, thereby giving logic a new expressiveness and power.

Modal Logic

Modal logic studies the modes of truths and how they relate to reasoning. Modes of truth are the various ways that a proposition can be true or false. These include necessary truths, necessary falsities, contingent truths, and contingent falsities. There are also possible truths and falsities.

In logic, a circumstance is *possible*, so long as its description is not self-contradictory. In this way, *possible* is not the same as *probable*. Some circumstances are possible although unlikely; for example, a person winning a million dollars in the state lottery 3 days in a row is a possible outcome but not a probable one. However, some things are not even possible because their descriptions are self-contradictory; for example, there is a woman who is taller than herself.

Aristotle founded modal logic, but the first modern system is ascribed to Clarence Irving Lewis, usually cited as C. I. Lewis. Today, work in modal logic rests on one or another version of *possible world semantics*. The notion of a possible world, of a way things might have or could have been, can be used to express the whole of logical theory and resolve theoretical problems that would otherwise remain unresolved. Modal logic is of interest to philosophers because many

of the interesting arguments in the history of philosophy—those about existence, the soul, free will, God, and so on are modal in nature. *The Nature of Necessity*, published in 1974, was one of the most important works in the history of modal logic. Alvin Plantinga presented a new version of Anselm of Canterbury's (1033–1109 C.E.) famous ontological argument for the existence of God by using modal logic, establishing Anselm's argument as valid, and defending it against objections.

Informal Logic

Whereas formal logic applies to areas of study that consider reasoning as abstract from other areas, informal logic is concerned with the informal aspects of logical theory. There are three important branches of informal logic: the study of definitions, the study of informal fallacies, and the study of inductive reasoning.

Definitions are explanations of the meanings of words or phrases. If we do not define words properly, our reasoning may become sidetracked, so the art of definition is an important element in logical theory.

A *fallacy* is an error in reasoning. The ability to recognize a fallacy is considered an important logical skill. Fallacies should not be persuasive, but they often are. They can be created unintentionally or intentionally to deceive people. Most fallacies involve arguments, whereas others can be involved in explanations, definitions, or other forms of reasoning. The term *fallacy* can be used more broadly to highlight false beliefs or causes of false beliefs. An informal fallacy is fallacious due to both its form and content, whereas formal fallacies are fallacious because of their logical form.

Aristotle was again the first person known to consider fallacies, and he produced a list of common informal fallacies—there are now a huge number catalogued.

Inductive reasoning is used in everyday life, in science and in the law as well as in numerous other contexts. There are 12 basic types of inductive reasoning, and understanding them can give us an insight into an extremely useful logical skill. During the 19th century, the English philosopher John Stuart Mill formulated the rules of induction

that remain in use today. Scientists run tests in laboratories aimed at getting to the causes of things, and health departments use the principles laid down by Mill, known as *Mill's rules* when the researchers are attempting to track down the sources of disease.

Other Types of Logic

Paraconsistent logic is a branch of logic which is nonexplosive, whereby a contradiction does not imply the truth of each and every sentence. Its rules of inference are such that not every statement follows from a contradiction. Paraconsistent logic gives up the classical *principle of explosion*, arguing that it should be possible to reason in a logical discerning way beginning with a contradictory theory and not ending in meaningless triviality.

Dialetheism claims that some contradictions, at least, are true. This nonformal system of logic states that were one to exist, this would be a *true contradiction*. For at least one sentence *S*, both *S* and $\neg S$ are true at the same time, rejecting the *law of contradiction* principle; in other words, rejecting the principle that lays the foundation for the entire realm of classical logic.

The Value of Logic

The principles of logic are not just an academic exercise; they are practical guides that can be used to correct reasoning. Just as the principles of accountancy ensure we keep a correct set of books and the principles of arithmetic aid us in adding, subtracting, multiplying, and dividing correctly, the principles of physics enable us to plan landings on the moon, the rules of photography help us to take better pictures, and so on. Not all of us reason correctly all the time; we tend to make logical errors at least occasionally. In paying attention to the principles of logic, we have the tools to work things out when matters get complicated.

Thora L. Fitzpatrick

See also Formal Sciences; Logic, Formal and Informal; Logic, Inductive

Further Readings

- Barnes, J. (1981). *Proof and the syllogism*. Oxford University Press.
- Bassham, G., Irwin, W., Nardone, H., & Wallace, J. (2004). *Critical thinking: A student's introduction*. McGraw-Hill.
- Govier, T. (2014). *A practical study of argument* (7th ed.). Cengage.
- Haskell, C. (2011). *Foundations of mathematical logic*. Dover.
- Herrick, P. (2011a). *The basic concepts of logical theory*. *The many worlds of logic*. <http://www.manyworldsoflogic.com/basicConcepts.html>
- Herrick, P. (2011b). *A Cook's tour of the fundamental ideas of logical theory*. *The many worlds of logic*. Author. <http://www.manyworldsoflogic.com/whatIsLogic.html>
- Lear, J. (1980). *Aristotle and logical theory*. Cambridge University Press.
- Mill, J. S. (2002). *A system of logic*. University Press of the Pacific.
- Oxford University. (n.d.). *Introduction to logic*. Author. <http://logic.philosophy.ox.ac.uk/>
- Plantinga, A. (1974). *The nature of necessity*. Oxford University Press.

The SAGE Encyclopedia of
**THEORY IN SCIENCE,
TECHNOLOGY,
ENGINEERING, AND
MATHEMATICS**

The SAGE Encyclopedia of
**THEORY IN SCIENCE,
TECHNOLOGY,
ENGINEERING, AND
MATHEMATICS**

2

Edited by

James Mattingly

Georgetown University

 **SAGE** reference

Los Angeles | London | New Delhi | Singapore | Washington DC | Melbourne



FOR INFORMATION:

SAGE Publications, Inc.
2455 Teller Road
Thousand Oaks, California 91320
E-mail: order@sagepub.com

SAGE Publications Ltd.
1 Oliver's Yard
55 City Road
London, EC1Y 1SP
United Kingdom

SAGE Publications India Pvt. Ltd.
B 1/I 1 Mohan Cooperative Industrial Area
Mathura Road, New Delhi 110 044
India

SAGE Publications Asia-Pacific Pte. Ltd.
18 Cross Street #10-10/11/12
China Square Central
Singapore 048423

Copyright © 2023 by SAGE Publications, Inc.

All rights reserved. Except as permitted by U.S. copyright law, no part of this work may be reproduced or distributed in any form or by any means, or stored in a database or retrieval system, without permission in writing from the publisher.

All third party trademarks referenced or depicted herein are included solely for the purpose of illustration and are the property of their respective owners. Reference to these trademarks in no way indicates any relationship with, or endorsement by, the trademark owner.

Printed in the United States of America

ISBN: 9781483347721

Acquisitions Editor: Kaitlin Ciarniello
Development Editor: Sanford Robinson
Production Editor: Astha Jaiswal
Copy Editor: Hurix Digital
Typesetter: Hurix Digital
Proofreaders: Dennis Web, Ellen Brink, Theresa Kay
and Wendy Jo Dymond
Indexer: Sheila Hill
Cover Designer: Candice Harman
Marketing Manager: Gabrielle Perretta

This book is printed on acid-free paper.

22 23 24 25 26 10 9 8 7 6 5 4 3 2 1

Contents

Volume 2

List of Entries *vii*

Reader's Guide *xi*

Entries

M	501	T	893
N	603	U	931
p	629	V	935
R	767	W	945
S	799		

Appendix: Resource Guide 957

Index 961

List of Entries

Abduction
Abstract Knowledge
Abstraction
Accuracy
Actor-Network Theory
Ad Hoc Hypothesis
Analysis
Applicability of Mathematics in Science
Artificial Intelligence
Assumption
Astronomy. *See* Cosmology
Authority
Axiom Schema
Axiomatic Theory

Belief Revision
Big Data
Biochemistry, 19th Century
Biochemistry, 20th Century
Biochemistry, Contemporary
Bioinformatics
Biology, Evolutionary
Biophysics, 19th Century
Biophysics, 20th Century
Biophysics, Contemporary
Biopsychosocial Model
Biostatistics

Category Theory
Cell Theory
Chemistry
Church–Turing Thesis
Climate Science
Cognitive Science
Communication Theory, Technical Overview
Complementary Medicine
Completeness
Complex Systems
Computer Science. *See* Cybernetics, 20th Century;
Software Engineering; Big Data

Concept
Conceptual Analysis
Conceptual Blending
Consilience
Cosmology
Cybernetics, 20th Century

Data and Phenomena
Data Mining. *See* Big Data
Data Models
Deduction
Developmental Systems Theory

Ecology. *See* Environmental Studies; Natural
Resources
Empiricism. *See* Epistemology; Evidence;
Experiment, Theory of
Engineering, Contemporary
Engineering, History of
Environmental Studies
Epistemology
Evidence
Evolutionary Psychology
Experiment, Theory of
Explanation

Fact Versus Theory
Falsifiability
Formal Sciences
Framework

Game Theory
Generalization
Genetic Drift
Geology
Geometry, Classical
Geometry, Non-Euclidean
Geophysics. *See* Geology
Germ Theory
Gödel’s Incompleteness Theorems

Health Care Science
Hypothesis Testing
Hypothetico-Deductivism

Induction
Infectious Disease Studies
Inference
Inference to the Best Explanation
Inferentialism
Informatics
Information Theory
Information Theory, Historical Background
Instrumentalism
Interpretation
Intuition
Intuitionism in Logic and Mathematics

Justification

Kinetic (Molecular) Theory
Knowledge

Laws, Scientific
Laws of Nature
Life Sciences, Contemporary
Linguistic Frameworks
Linguistics, Contemporary
Linguistics, Historical
Logic, Formal and Informal
Logic, Inductive
Logic and Language
Logical Theory

Mathematics, Antiquity
Mathematics, Enlightenment
Mathematics, Middle Ages
Mathematics, 19th Century
Mathematics, Renaissance
Mathematics, 20th Century
Measurement
Medicine, Antiquity
Medicine, Contemporary
Medicine, Enlightenment
Medicine, Middle Ages
Medicine, 19th Century
Medicine, Renaissance
Medicine, 20th Century. *See* Medicine,
Contemporary; Medicine, 19th Century
Mental Models
Metaphysics

Metatheory
Meteorology. *See* Climate Science
Modeling

Nanotechnology
Natural Resources
Natural Selection
Neuroscience
Number Theory. *See* Mathematics, 19th Century;
Mathematics, 20th Century

Paraconsistency
Paradigm
Paradoxes
Perturbation Theory
Phenomenalism
Philosophy of Mind
Philosophy of Science
Physical Theory
Physics, Antiquity
Physics, Contemporary
Physics, Enlightenment
Physics, Gravitation Theory
Physics, Middle Ages
Physics, 19th Century
Physics, Quantum Theory
Physics, Renaissance
Physics, Solid-State
Physics, Thermodynamics
Physics, 20th Century
Plate Tectonics
Pragmatism
Prediction
Pseudoscience
Punctuated Equilibrium

Rational Mechanics
Rationality
Realism in Mathematics
Realism in Science. *See* Scientific
Realism
Reasoning
Received View of Theories. *See* Theory,
Syntactic Conception
Refutation
Relative Consistency

Scientific Realism
Scientific Revolutions
Semantics (Introduction to Theory)

Semantics (Scientific and Empirical)	Taxonomy
Set Theory	Theory Change
Social Construction of Scientific Knowledge	Theory Construction
Software Engineering	Theory Structure
Soundness	Theories, Semantic Conception of
Speculation	Theories, Syntactic Conception of
Statistical Inference, Bayesian	Thought Experiments, Scientific and Philosophical
Statistical Inference, Frequentist	Truth
Statistics	Understanding
Statistics, Completeness in	Values in Science
Structuralism (Introduction to Theory)	Vitalism
Syntax (Introduction to Theory)	Warrant
Syntax (Scientific and Empirical)	Worldview
Systems Science	

Reader's Guide

General Theory of Science

Abduction
Abstract Knowledge
Abstraction
Accuracy
Ad Hoc Hypothesis
Analysis
Authority
Belief Revision
Church–Turing Thesis
Communication Theory, Technical Overview
Completeness
Concept
Conceptual Analysis
Conceptual Blending
Consilience
Epistemology
Evidence
Experiment, Theory of
Explanation
Fact Versus Theory
Falsifiability
Formal Sciences
Generalization
Hypothesis Testing
Hypothetico-Deductivism
Induction
Inference
Inference to the Best Explanation
Inferentialism
Instrumentalism
Interpretation
Intuition
Justification
Knowledge
Mental Models
Metaphysics
Metatheory
Modeling
Paraconsistency

Paradigm
Paradoxes
Phenomenalism
Philosophy of Mind
Philosophy of Science
Physical Theory
Pragmatism
Prediction
Pseudoscience
Rationality
Realism in Mathematics
Reasoning
Scientific Realism
Scientific Revolutions
Social Construction of Scientific Knowledge
Speculation
Statistics
Taxonomy
Theory Change
Thought Experiments, Scientific and Philosophical
Truth
Understanding
Values in Science

Nature and Structure of Theories

Axiom Schema
Axiomatic Theory
Data Models
Framework
Geometry, Classical
Geometry, Non-Euclidean
Gödel's Incompleteness Theorems
Instrumentalism
Intuitionism in Logic and Mathematics
Laws, Scientific
Laws of Nature
Linguistic Frameworks
Modeling
Paradigm
Philosophy of Science

Rational Mechanics
 Received View of Theories
 Relative Consistency
 Set Theory
 Theories, Semantic Conception of
 Theories, Syntactic Conception of
 Theory Construction
 Theory Structure

Formal Disciplines

Formal Sciences
 Game Theory

Computer Science

Artificial Intelligence
 Big Data
 Bioinformatics
 Biostatistics
 Church–Turing Thesis
 Complex Systems
 Cybernetics, 20th Century
 Informatics
 Information Theory
 Information Theory, Historical Background
 Software Engineering
 Statistical Inference, Bayesian
 Statistical Inference, Frequentist
 Statistics
 Statistics, Completeness in
 Systems Science

Logic and Mathematics

Abduction
 Abstraction
 Analysis
 Axiom Schema
 Axiomatic Theory
 Category Theory
 Church–Turing Thesis
 Communication Theory, Technical
 Overview
 Concept
 Cybernetics, 20th Century
 Deduction
 Generalization
 Geometry, Classical
 Geometry, Non-Euclidean
 Gödel's Incompleteness Theorems
 Induction
 Inference
 Intuitionism in Logic and Mathematics

Justification
 Knowledge
 Linguistics, Contemporary
 Linguistics, Historical
 Logic, Formal and Informal
 Logic, Inductive
 Logic and Language
 Logical Theory
 Mathematics, 19th Century
 Mathematics, 20th Century
 Mathematics, Antiquity
 Mathematics, Enlightenment
 Mathematics, Middle Ages
 Mathematics, Renaissance
 Paraconsistency
 Paradoxes
 Perturbation Theory
 Rational Mechanics
 Realism in Mathematics
 Reasoning
 Refutation
 Relative Consistency
 Semantics (Introduction to Theory)
 Semantics (Scientific and Empirical)
 Set Theory
 Soundness
 Statistics
 Syntax (Introduction to Theory)
 Syntax (Scientific and Empirical)
 Understanding

Empirical Disciplines

Physical Theory

Biological Science

Biochemistry, 19th Century
 Biochemistry, 20th Century
 Biochemistry, Contemporary
 Bioinformatics
 Biology, Evolutionary
 Biophysics, 19th Century
 Biophysics, 20th Century
 Biophysics, Contemporary
 Biostatistics
 Cell Theory
 Complementary Medicine
 Developmental Systems Theory
 Environmental Studies
 Evolutionary Psychology
 Game Theory
 Genetic Drift

Germ Theory
 Health Care Science
 Infectious Disease Studies
 Life Sciences, Contemporary
 Medicine, 19th Century
 Medicine, 20th Century
 Medicine, Antiquity
 Medicine, Contemporary
 Medicine, Enlightenment
 Medicine, Middle Ages
 Medicine, Renaissance
 Natural Selection
 Neuroscience
 Punctuated Equilibrium
 Vitalism

Chemistry

Biochemistry, 19th Century
 Biochemistry, 20th Century
 Biochemistry, Contemporary
 Chemistry

Cognitive Sciences

Artificial Intelligence
 Belief Revision
 Big Data
 Biopsychosocial Model
 Cognitive Science
 Conceptual Blending
 Evolutionary Psychology
 Linguistic Frameworks
 Linguistics, Contemporary
 Linguistics, Historical
 Mental Models
 Neuroscience
 Philosophy of Mind
 Semantics (Introduction to Theory)
 Semantics (Scientific and Empirical)
 Syntax (Introduction to Theory)
 Syntax (Scientific and Empirical)

Earth and Space Sciences

Climate Science
 Cosmology
 Environmental Studies
 Geology
 Natural Resources
 Plate Tectonics

Engineering

Big Data

Bioinformatics
 Communication Theory, Technical
 Overview
 Complex Systems
 Cybernetics, 20th Century
 Engineering, Contemporary
 Engineering, History of
 Informatics
 Information Theory
 Modeling
 Perturbation Theory
 Software Engineering
 Statistics

Physics

Biophysics, 19th Century
 Biophysics, 20th Century
 Biophysics, Contemporary
 Complex Systems
 Cosmology
 Environmental Studies
 Geophysics
 Information Theory
 Kinetic (Molecular) Theory
 Physical Theory
 Physics, 19th Century
 Physics, 20th Century
 Physics, Antiquity
 Physics, Contemporary
 Physics, Enlightenment
 Physics, Gravitational Theory
 Physics, Middle Ages
 Physics, Quantum Theory
 Physics, Renaissance
 Physics, Solid-State
 Physics, Thermodynamics
 Plate Tectonics
 Rational Mechanics
 Systems Science

Technology

Artificial Intelligence
 Cybernetics, 20th Century
 Engineering, Contemporary
 Engineering, History of Health Care
 Science
 Informatics
 Modeling
 Nanotechnology
 Software Engineering

M

MATHEMATICS, ANTIQUITY

The period of antiquity usually refers to the time the first records of human civilization were recorded up to the collapse of the Western Roman Empire. In reality, this temporal reference point is arbitrary because mathematical activities did not exist only in this part of the world. The adoption of this time framework was influenced by the state of mathematical practices today, which is profoundly Greek axiomatic-deductive in structure, and by the Western points of view of historians of mathematics. However, mathematics has always been, and still is, a social activity that societies have practiced—one which is linked to their ways and cultures. Historically, it emerged in different ways. Mathematizing was brought about by practical necessities such as agriculture, architecture, and the scientific activities of those times. In many instances, these practical techniques led to more theoretical inquiries such as those exploring the nature of shapes and measurement.

Through the study of the history of mathematics, it can be seen that this body of knowledge was transmitted and transformed, rediscovered and synthesized in different kingdoms, states, epochs, and cultures. The study of mathematics in antiquity is the study of the interweaved nature of this most basic human activity.

Mathematics cannot be detached from its context. It is important to understand that these activities were part of societies and cultures, each

with their own traditions, beliefs, and needs. Therefore, it is natural that mathematics developed differently in different parts of the world. In interpreting history, it is important to define which activities could be categorized as mathematics. For practicality, many of today's historians have simply considered activities that appeared similar to our modern mathematics, like counting, arithmetic, measurement, and geometry. In the past, however, mathematics was more intertwined with religion, science, commerce, governance, engineering, and architectural pursuits. This is very different from the current state of mathematics, which despite being still highly influential to these same fields, is regarded as its own separate area of intellectual inquiry.

The Importance of Historical Artifacts

The most important factor that drives our knowledge about mathematics in antiquity is the availability of artifacts. Our knowledge of the ancient world is limited because only a few documents have survived. The consequence is that the mathematical histories of societies with fewer archaeological materials might appear to have engaged less in mathematical activities. In addition, societies that were relatively small would have fewer artifacts likely to have been preserved. This is why in the study of ancient mathematics, the focus is usually on the Egyptians, the Mesopotamians, the Chinese, and the Greeks. Larger societies produced more mathematical materials that lasted

much longer; yet even among those societies, modern understanding varies because of geography and the ways they recorded their mathematical activities. For example, Mesopotamians used clay, a more durable material than the papyri used by Egyptians, which disintegrated easily. This is why more is known about Mesopotamian mathematics than Egyptian. For Chinese mathematics, records from this time were almost surely written for the central government, which could have resulted in a skewed view of the purposes of mathematics. Documents written by Chinese mathematics practitioners who were not in civil service are rare. Finally, Greek mathematics was preserved through documents copied by the medieval Persians and Arabs, who translated and transformed them, later transmitting them back to Europe. Without these documents, significantly less would have been known. It is therefore important to remember that our understanding can never be complete because we are basing our findings only on limited archeological artifacts.

Practical and Theoretical Mathematics

Current historians of mathematics recognize that interpretations could be misguided by presentism. The most fundamental aspect of mathematical practice today is the concept of proofs, generalized for all possible situations, using an axiomatic-deductive style. A proof starts with axioms, or assumptions that are considered true, and then is followed by an unbroken sequence of argumentation that arrives at a conclusion, the statement that needs to be proven. This replicates the style found in Euclid of Alexandria's textbook *The Elements*. This ancient Greek way of mathematics is the mold that modern mathematicians aspire to follow. Exactness is a necessity and approximations are not valid. Many mathematical problems are not directly integrated with practical activities and many are not even concerned with things that are tangible. Mathematical text is conceptually dense. It assumes the reader knows the previously proven theorems necessary to understand the text. These are characteristics of the practices of small groups of mathematicians from the later period of ancient Greek society before it became part of the Roman Empire. The leaning toward the more *theoretical* or not directly useful as opposed to

practical mathematics is what modern mathematicians have adopted and was the practice copied from this group of ancient practitioners. Nevertheless, the majority did not share this type of intellectual endeavor. This style of mathematics was not the only prevailing practice, nor was it the most common in this period.

Categorization by Geography

In the ancient world, mathematical ideas were exchanged across societal borders. The lifetimes of civilizations overlapped and influenced each other's mathematical progress. Despite this, historians usually adopt a geographical demarcation rather than a temporal framework for their analyses. Consequently, the history of ancient mathematics is not subdivided such that mathematical activities for a certain century include all societies creating mathematics at that time. As noted earlier, this is because of the limited archeological artifacts and knowledge available. For practicality, historians usually use geographical divisions including Egyptian, Mesopotamian, Chinese, and Greek mathematics.

Egyptian

The history of ancient Egypt is typically divided into three phases: the Old Kingdom (1550–1069 B.C.E.), the Middle Kingdom (2055–1650 B.C.E.), and the New Kingdom (1550–1069 B.C.E.) when only one king ruled the land. Culture thrived and expanded during these periods because of political stability. After these kingdoms, Egypt was conquered by Alexander the Great and ruled throughout the Hellenistic period (305–30 B.C.E.) by various Ptolemaic dynasties (Macedonian Greek royal families) until it eventually became a province of the Roman Empire.

Current knowledge of Egyptian mathematics is limited because artifacts available for study are scarce. Only a few have been found containing mathematical text including the Rhind, Kahun, and Berlin papyri, the Akhmim wooden tablets, a leather roll, and the Moscow papyrus. All of them date to the Middle Kingdom. The reason for the rarity of artifacts is because excavations are limited to uninhabited areas close to the Nile valley, desert areas where papyri would have a greater chance of survival. Papyri would have less chance

of being preserved in the more humid inhabited areas adjacent to the river Nile. This circumstance makes it impossible to assess the extent of Egyptian mathematical activity at that time.

A fully developed decimal system of numeration already existed even before Egypt was unified under one ruler. In terms of writing, *hieroglyphs*, meaning sacred engraved text, were used for inscriptions on stone monuments, while a cursive form of writing called *hieratic* was used on papyrus for everyday purposes. *Demotic* is a later and more cursive version of hieratic. Only hieratic and demotic were used for mathematical text. The motivation for doing mathematics was practical; for instance, various technological pursuits. Discoveries of school texts used to train scribes describe problems in measuring areas of land, calculating agricultural production, and calculating the slopes of pyramids. Units of measurement were based on body parts. For example, a cubit corresponds to the forearm, which is 7 *palms* long, a unit that in turn is equivalent to 4 *fingers*. To perform fast calculations, Egyptians practiced multiplication and division through successive doubling or halving, working with powers of 10, using even numbers, and powers of two and their reciprocals. They used a system of fractions that is different from modern practices but made the magnitude immediately apparent: for example, expressing $5/19$ as $(1/6, 1/12, 1/114, 1/228)$. They also used precalculated tables to aid in division problems.

Important mathematical activities were used for various engineering projects such as stone mining. The calculation of stone volumes was essential to determine how many men could carry the weight and how long the work would take. These activities required people to have a simple way of calculation, a type of mathematics that can be understood by everyone involved. The Rhind papyrus contains a problem regarding the calculation of *seked*, a concept equivalent to the modern-day concept of slope. Mathematics was also important for agriculture, especially measuring land under cultivation. Every year, the Nile would flood the plains and the level of the flood would indicate the agricultural output for the next year. Therefore, it was important to measure and record these levels. In addition, the boundaries of agricultural lands that were previously flooded had to be reestablished as soon as the water receded. Finally,

mathematics was also used to calculate quantities of grain, the ratio of grains to output for bread and beer known as the *pefsu*, forecasting productivity, and calculating taxes.

The practicality of mathematics found on currently existing papyri, sources that were used in educating scribes, might lead to a premature conclusion that this is the only type of mathematics in ancient Egypt. It is important to note that while Egyptian mathematical activities were highly practical, some recorded problems were not. For instance, calculating circular pieces of land, while outside practical applications, were extensions of more common land-measuring problems. In this case, Egyptians did not use the concept of *pi* to calculate the areas of circles but used a formula to give sufficient approximations. What they did was to subtract $1/9$ of the diameter from itself and square the answer, resulting in a good approximation of the area. This is an example of practical problems giving rise to more theoretical investigations.

Mesopotamian

This civilization is also referred to as *Babylonian*, although the city of Babylon was not always the center of culture in this region. This area was called *Mesopotamia*, meaning country between two rivers, because of the rivers Tigris and Euphrates that bordered it. Owing to this water source, the area became viable for agriculture and it earned the name Fertile Crescent. It also was open to invasions and had multiple rulers from different regions. Unlike the Egyptian civilization, the Mesopotamian area was composed of various independent city-states. The cuneiform or wedge-shaped writing system that developed from this area possibly antedates the Egyptian hieroglyphs and may have been the earliest form of written communication. These were impressions made on clay tablets, making them more durable than the Egyptian papyri.

Unlike Egyptian mathematics, there are more artifacts with mathematical text that have survived and have been excavated from this region. Through the study of ancient problems, it becomes evident that ancient Mesopotamian scribes used their mathematical knowledge in various areas such as land measurement, business calculations such as silver and grain exchanges, and agriculture. There also are examples that demonstrate mathematical

inquiries and abstractions beyond everyday specific practical applications. For example, the Mesopotamians knew how to solve three-term quadratic equations. What is called now as the Pythagorean theorem was also widely used. It is unknown whether these algebraic activities were for practical purposes or for general understanding, but many historians believe there is justification to say that they went beyond the practical. There might be awareness of these concepts hidden in specific practical problems. For instance, in some word problems, concepts of length and width might have been interpreted as variables, equivalent to the x and y symbols used in modern mathematics.

The numeration system from this region was sexagesimal. This might have evolved because it was easier to do calculations with it. The sexagesimal numbers are easier to divide by two, three, four, five, six, 10, 12, 15, 20, and 30, making it convenient for commercial-related arithmetic. The system used was also positional, which means the placement of the number signifies that it is a multiple of different orders of magnitude. For example, in our current decimal system, the number 532 signifies that the 5 is 5 multiplied by 100, the 3 means 3 multiplied by 10, and 2 by 1. Although initially Mesopotamians did not have a symbol for zero and just left a blank for the empty spot, a symbol eventually was devised, consisting of two small wedges placed diagonally. They were able to expand this positional property for fractions; therefore, they had computational power similar to that of our modern decimal fractional notation. This made it easier to get more accurate approximations, such as for calculating the value of the square root of two. They had an algorithm for calculating any square root, which later was attributed to others, including the Greek scholar Archytas, Heron of Alexandria, and Isaac Newton. Their computational sophistication is also evident in the precalculated tables that were used by scribes, and which contained successive powers of a given number, almost like modern logarithmic tables; these, however, were used for specific problems and not for general calculations.

Chinese

Some historians believe that the Chinese developed their mathematics independently from other

civilizations in the ancient world. Before the forced unification of the region by the Qin dynasty in 221 B.C.E., there were many master scholar–student relationships that flourished, making possible the transmission of mathematical knowledge. After this period, the Chinese central government started employing intellectuals to serve it directly. Most intellectuals, including mathematics practitioners, were government workers. These scholars supported the government ideologically as both teachers and writers. Eventually, the Han dynasty replaced the Qin and established the centralized civil service examination system, still enlisting the most capable and literate men to be part of the government. With this process, the government controlled the ideologies that prevailed at that time. The dynasty privileged Confucian thought over all other schools and made this the official state ideology.

One of the necessary forms of expertise for a civil servant was the *suan*. Literally, this referred to counting rods used for arithmetic, but it also meant the practice of using them. The underlying system was decimal. *Suan* was not the most important skill government employees had, nor was it a separate area of intellectual inquiry. The Chinese at that time did not differentiate between the study of nature and the study of mathematics, and they investigated the connections between mathematics and areas such as physics, cosmology, astronomy, and music.

Two major treatises were written in ancient China, the *Zhou bi suan jing*, translated to “Arithmetic classic of the gnomon of Zhou” and the *Jiu zhang suan shu* or the “Nine chapters on mathematical procedures.” The later document contained 246 problems on surveying, business, engineering, agriculture, solutions of equations, and right triangles. It followed the problem-and-answer method also found in other ancient texts from other regions of the world. The method for calculating the areas of basic shapes such as triangles, circles, rectangles, and trapezoids were also discussed in this document. A significant section of the *Nine chapters* was on solutions to problems on simultaneous linear equations using both positive and negative numbers, something practitioners were accustomed to because they used two sets of rods, red and black, to denote negative and positive.

During the Han dynasty, *suan* remained an unimportant subject among the other fields of study. In fact, the two treatises written about it were categorized under astronomical texts. Mathematical inquiry, or creating a new method of *suan*, was not an important endeavor at that time. This was until the period of division, from 221 to 581 C.E., when the empire was broken apart into competing kingdoms. Suddenly, the ideologies from the past were rediscovered, some dealing with logic and mathematical concepts. This was the period when an intellectual named Liu Hui wrote a commentary about the *Nine chapters*, describing how it actually worked and adding discussions such as the correct calculation for the circumference of the circle. This reworking of the *Nine chapters* was a start of the re-blossoming of mathematics in ancient China.

Greek

The unique mathematical style that developed in the ancient Greek world was, as noted earlier in this entry, the axiomatic-deductive system, a way of mathematizing utilized in major ancient mathematics text such as *The Elements* by Euclid. This was not the mathematics practiced by the majority but only by a social elite who were more interested in theoretical matters—those mathematical ideas far removed from practical applications. Nonetheless, this is the approach adopted by modern mathematicians and the inherent structure of current mathematical writing. The Euclidian style was absent in earlier times and mathematical writings from that period implied that the reader was meant to understand abstract methods by dealing with actual cases. The practical mathematics problems were recipe-like, and rhetorical structures were evident. In the earlier periods of ancient Greek civilization and among practical mathematics practitioners, the concepts of definitions, proofs, and arguments were not known. Greek theoretical mathematics, which was exclusively geometric, only appeared toward the end of the 5th century B.C.E.

The common person's practical mathematics was the basis from which theoretical mathematics evolved over centuries. Euclidian-style mathematics appeared only in the later periods of Greek

civilization. Pebble arithmetic, similar to techniques using an abacus, was offered by practitioners-for-hire for various types of calculations. Not much is known about these professionals, but it is known that some Pythagoreans in 5th century B.C.E. Greece used their knowledge for semi-religious practices. In this period, there were also geometers who were hired to measure areas and volumes.

The Greeks adopted the literary and mathematical traditions of civilizations in the same region. Arithmetic knowledge came from Old Babylon, and fractional computation and geometry from Egypt. In fact, many ancient Greek writers such as Socrates and Aristotle attribute the beginnings of mathematics to the Egyptians. Aside from Mesopotamian and Egyptian mathematics, they also were able to transform the Phoenician writing system into their own. For example, Thales of Miletus (ca. 626–ca. 548 B.C.E.), a mathematician who made many discoveries, actually had ideas already known to Babylonians a millennium before. Another popular ancient mathematician, Pythagoras of Samos (b. 569 B.C.E.), traveled to Egypt, Babylon, and possibly India, where he learned various mathematical and religious ideas. Not surprisingly, mathematics and religion were intertwined within the Pythagorean school of thought. Number mysticism was an integral part of the Pythagoreans' beliefs, and numbers were elevated to a higher status. It was said that under them, the concepts of theoretical and practical mathematics, that which is for training the mind and that for real-life applications, were demarcated. These ideas were later promulgated by Plato (ca. 429–347 B.C.E.) when he distinguished between arithmetic as number theory for philosophers and logistic for businessmen and men of war.

The center of Greek theoretical mathematics was Athens for more than a century, until it moved to Alexandria in Egypt. In this city during the 5th century B.C.E., many mathematicians contributed profoundly to the shape of mathematical practice. Some of the important problems tackled during this period were the squaring of the circle, the duplication of the cube, paradoxes on motion, and ideas about infinity. Most of the mathematicians at this time were associated with the Academy of Plato in Athens. However, in 323 B.C.E., Alexander the Great died and the empire collapsed, leading to

the territories being divided and the center of mathematics moving to Alexandria. This was where intellectuals like Euclid, Archimedes, and Apollonius left their legacies that led to the more advances in mathematics. It developed in this period until it declined during under the Roman Empire, which was less hospitable to scholars, eventually moving its center to the Byzantine Empire.

Mark Causapin

See also Axiomatic Theory; Geometry, Classical

Further Readings

- Boyer, C. B. (1968). *A history of mathematics*. Wiley.
- Mann, T. (2011). History of mathematics and history of science. *Isis*, 102(3), 518–526.
- Robson, E., & Stedall, J. (Eds.). (2008). *The Oxford handbook of the history of mathematics*. Oxford University Press.
- Von, Staden. H. (1992). Affinities and elisions: Helen and Hellenocentrism. *Isis*, 83(4), 578–595. <https://doi.org/10.1086/356290>

MATHEMATICS, ENLIGHTENMENT

The Age of Enlightenment describes a philosophical and intellectual movement that took place during the 17th and 18th centuries. During this time, there occurred a shift in thinking toward reason, logic, and freedom, views that were supported, for example, by the rediscovery of earlier, mathematical texts. Logic, although not a new invention, having been much in use by the ancient Greeks, was now included in a worldview which argued that empirical observation could be used to reveal the true nature of the universe and the self. Enlightenment thinkers held that science could explain humankind, and that humanity's history had been marked by progress, a progress that could be continued if rational methods were followed. Enlightenment thinkers also believed that human character could be improved through education and reason, and that the universe was a functioning machine with the capacity for alteration. Enlightenment thinkers wanted not only to understand the world but also to change it for the better using reason and science.

Although there is no definitive starting point or ending point for the Enlightenment, it is generally supposed to be a 17th- and 18th-century phenomenon. Some historians give the English Civil Wars as key start dates for the philosophies of Enlightenment figures including the philosopher Thomas Hobbes, with the close of the era often attributed either to the death of Voltaire in 1778, or the beginning of the French Revolution in 1789.

The Enlightenment, coming as it did in the wake of the Renaissance, saw an unprecedented explosion in mathematical and scientific ideas across Europe. This is why the Enlightenment was also called the *Age of Reason*. Mathematics followed hard on the heels of the Copernican Revolution of the 16th century as well as the revolutionary discoveries made by Galileo, Tycho Brahe, and Johannes Kepler, Kepler's exploration of the solar system having led to Kepler's mathematical laws of planetary motion. The invention of the logarithm by John Napier early in the 17th century made a big contribution to the disciplines of astronomy and mathematics by rendering some formerly complex calculations far simpler. In fact, this was one of the most significant mathematical developments of the time, paving the way for the innovations of Kepler and Newton. French astronomer and mathematician Pierre Simon Laplace working much later, in the 19th century, noted that Napier had halved the work of astronomers with his groundbreaking developments.

Early Enlightenment Mathematics

René Descartes (1596–1650) is often considered a leading figure of the enlightenment. A metaphysician, mathematician, and natural philosopher, he was not only responsible for changing the course of philosophy but is also considered one of the earliest Enlightenment mathematicians, his main contributions being analytical geometry and the theory of vortices. For Descartes, analytical geometry is not simply concerned with the application of algebra to geometry, as set forward by Archimedes and others. Descartes's form of analytic (also called *Cartesian*) geometry uses algebra to describe geometry. His method of representing unknowns in equations using x , y , and z and *knowns* as a , b , and c was to become the norm. He also pioneered the standard notation that makes

use of superscripts to illustrate the powers or exponents, for example, in squared numbers. He was the first mathematician to assign algebra a fundamental place in our system of knowledge and believed that algebra's function was to automate and mechanize reasoning, in particular reasoning about unknown or abstract quantities.

Descartes's understanding of algebra was profound. He proposed that the number of distinct roots of an equation is equal to the degree of the equation. He was also willing to consider negative (he termed them false roots) and imaginary roots. He developed a rule for determining the number of positive and negative roots in an equation, which became known as *The Rule of Descartes*. The rule states that an equation can have as many positive roots as it contains changes of sign, from + to – or from – to +; and as many negative roots as the number of times two + signs or two – signs are found in succession.

In *La Geometrie* (1637), Descartes outlined his revolutionary method for geometrical problem-solving, referred to as his *geometrical calculus*. He offered innovative algebraic techniques for analyzing geometrical problems and novel ways of understanding the connection between the construction of a curve and its algebraic equation. He also proposed an algebraic classification of curves based on the degree of the equations used in the representation of the curves. During Descartes's time, mathematicians were battling with questions surrounding geometrical proofs, especially with regard to the criteria for identifying curves. In respect of this *La Geometrie* had a huge impact.

Descartes established analytic geometry as a way of visualizing algebraic formulas. He developed a coordinate system as a way to locate points on a plane. The coordinate system comprises two perpendicular lines called *axes*. The vertical axis is designated as the *y* axis, and the horizontal axis is designated as the *x* axis. The intersection point of the two axes is called the *origin*, or point zero. The position of any point on the plane can be identified by locating how far, perpendicularly from each axis, the point lies. The position of the point in the coordinate system is specified by its *x* and *y* coordinates. This is written as (x, y) . The coordinate system is also known as the *Cartesian coordinate system*, after the Latin version of the name Descartes.

Descartes created the coordinate system to locate points on a plane, but the system later evolved into what we call *analytic geometry*. The basic principle of analytic geometry is that all pairs of values that satisfy the equation are coordinates of points on a curve. Conversely, all points on the curve have coordinates that satisfy the given equation *y*. Analytic geometry gives us the means to graphically express the relationship between two variables that are functionally related to each other. Described as a sort of *acid test* of the validity of a physical law or theorem, its coordinate system has had a huge impact, enabling analytic geometry to be applied to a wide variety of disciplines including physics, chemistry, economics, astronomy, and sociology. On this basis, Descartes has been described as having renovated a whole domain of human thought.

Foundational Mathematics

One of the most influential works of the early Enlightenment was Isaac Newton's *Principia Mathematica* (often referred to simply as the *Principia*). This book, published in 1687, is considered one of the most influential works in the history of science, and it would continue to dominate scientific views of the universe for centuries to come. Although it remains in the public consciousness largely owing to its famous explanation of the force of gravity—and to the apocryphal story of its author's having derived the principle when, sitting under an apple tree, he its fruit falling to the ground—Newton was a consummate mathematician and greatly influenced the path of mathematical development. He pioneered a new approach to mathematics, that of infinitesimal calculus, a calculus based on earlier work by mathematicians including Descartes, Pierre de Fermat, and fellow Englishmen Isaac Barrow and John Wallis. Unlike the static geometry of the Greeks, calculus enabled mathematicians to make sense of motion and dynamic change, including such things as the orbits of planets and the motion of fluids. Newton discovered that it was relatively easy to represent and calculate the average slope of a curve (e.g., the increasing speed of an object on a time–distance graph). However, since the slope of a curve constantly varies there was no way to identify the exact slope at any specific point on it—that is, to

work out the slope of a tangent line to the curve at that particular point. Using intuition, we can approximate the slope at a particular point by taking the average slope or *rise over the run* of increasingly smaller segments of the curve. When the segment of the curve under consideration approaches zero in size (where there is an infinitesimal change in x), the calculation of the slope gets nearer and nearer to the exact slope at a point.

Thus, Newton calculated a derivative function $f'(x)$ which gives the slope at any point of a function $f(x)$. This method of calculating the derivative of a curve or function is termed *differential calculus*, or differentiation. Having established the derivative function of a particular curve, it is relatively straightforward to calculate the slope at any particular point on that curve by inserting a value for x . The opposite of differentiation is integration, or *integral calculus*, and together these two perform the main operations of calculus. Newton's Theorem of Calculus states that differentiation and integration are inverse operations. If a function is firstly integrated and subsequently differentiated (or the other way around), then the original function will be retrieved.

Theory of Calculus

Gottfried Wilhelm Leibniz (1645–1716) also occupies a very important place in the history of mathematics. A contemporary of Newton, he developed a theory of calculus very similar to Newton's and in fact was the first to publish his work, which led to a certain amount of controversy over who was first to develop the theory. History in fact finally favored Leibniz's notation, which is still used by mathematics today. Leibniz was also involved in reviving the ancient method of matrices as a way of solving equations, which involved arranging linear equations in an array so they could be manipulated to find a solution.

This paved the way for the linear algebra of Carl Friedrich Gauss (1777–1855). He pioneered the development of the binary system and introduced notions of self-similarity, as well as the principle of continuity, foreshadowing the mathematical discipline of topology. Leibniz also worked on a practical calculating machine, which used the binary system, and could multiply, divide, and extract roots. This extraordinary

machine was the forerunner of today's computers. Leibniz is credited with developing the binary number system. He also worked extensively in the area of logic, even though he published nothing in his lifetime on the subject. In working drafts, we can see his setting out the principal properties of what are now known as *conjunction*, *disjunction*, *identity*, *negation*, *set inclusion*, and the *empty set*.

Toward the end of the 17th century, the focus shifted to the followers and disciples of Newton and Leibniz who applied their ideas on calculus to solving problems in astronomy, physics, and engineering. One name in particular stood out during this time, that of Bernoulli, the surname of two brothers, Jacob and Johann, who influenced several generations of outstanding mathematicians. They advanced Leibniz's theory of infinitesimal calculus, developing a calculus known as the *calculus of variations* by generalizing and extending it. The Bernoulli brothers also worked on Pascal and Fermat's earlier probability and number theory.

The Bernoullis lived in Basel, Switzerland, also hometown to influential Enlightenment mathematician Leonhard Euler. The brothers corresponded with Leibniz and were among the first mathematicians to study and comprehend infinitesimal calculus, as well as to apply it to new problems. They were also instrumental in helping to make calculus the cornerstone of mathematics, something which remains the same today. As well as being disciples of Leibniz, however, the Bernoullis made important contributions of their own. They concerned themselves with the problem of designing a sloping ramp, so that a ball could roll from top to bottom in the shortest time. Johann Bernoulli was able to demonstrate via calculus that neither a straight ramp nor a curved one with an extreme initial slope were the optimal answer. In fact, a less steeped curve, or kind of upside-down cycloid, would allow for the fastest descent of the ball. This application illustrated the use of the calculus of variations developed by the Bernoullis, which has been found useful in multiple areas including engineering, architecture, and finance. In 1713, Jacob Bernoulli's *The Art of Conjecture* was published, after his death. This consolidated existing probability theory and set out his theory of permutations and combinations, as well as some

aspects of Bernoulli number theory. He also produced papers on transcendental curves and was first to develop a technique for solving separable differential equations. Polar coordinates, which describe the location of points in space using distances and angles, are also included in his list of accomplishments, and he was the first mathematician to refer to the word *integral* as it related to the area under a curve.

Jacob Bernoulli was also responsible for discovering the approximate value of the irrational number e . He made this discovery while examining compound interest on loans. Compounded at 100% interest per year, £1 becomes £2 after 1 year. When compounded half-yearly, it produces £2.25; quarterly, £2.44; monthly £2.61; weekly £2.69; daily £2.71, and so on. If interest were to be compounded on a continual basis the £1 would tend toward a value of £2.7182818 after 1 year, and this value became known as e . In algebraic terms, this is the value of the infinite series $(1 + \frac{1}{1})^1 \cdot (1 + \frac{1}{2})^2 \cdot (1 + \frac{1}{3})^3 \cdot (1 + \frac{1}{4})^4 \dots$

Early 18th-Century Mathematics

Euler was a prolific mathematician of the renaissance. During his lifetime, he produced no less than 900 completed works, on average one mathematical paper a week. His interests in mathematics ranged from geometry and calculus to trigonometry, algebra and number theory. Much of the mathematical notation we use today originated with Euler, including e , i , $f(x)$, Σ and the use of a , b , and c as constants and x , y , and z as unknowns. He also attempted to standardize other symbols including π and the trigonometric functions. Euler's work helped to turn mathematics into an international subject which invited collaboration from mathematicians right across the world.

Euler was famous for producing what is considered one of the most beautiful mathematical equations of all time $e^{i\pi} = -1$, generally known as *Euler's Identity*. This equation combines calculus, arithmetic, complex analysis, and trigonometry into a remarkable mathematical formula. Another of his discoveries, Euler's Formula, also demonstrated the deep relationships between complex numbers, trigonometry and exponentials.

Euler's reputation was sealed by the discovery of the calculation of infinite sums in 1735. The

Bernoulli had tried but failed to solve this problem, so it had become known as the *Basel problem*. When asked what the precise sum of the reciprocals of the squares of all the natural numbers to infinity was, that is, $\frac{1}{1^2} + \frac{1}{2^2} + \frac{1}{3^2} + \frac{1}{4^2} \dots$ (a zeta function using a zeta constant of 2), Daniel Bernoulli stated that he believed it to be approximately $1\frac{3}{5}$. Euler's method, however resulted in an answer of $\pi^2/6$. He also demonstrated that the infinite series was equivalent to an infinite product of prime numbers. This identity was later to stimulate the investigation of complex zeta functions by Bernhard Riemann.

In the same year, 1735, Euler solved the logical and intransigent mathematical problem of the Seven Bridges of Königsberg. This enduring problem had remained unsolved for many years. The ancient city of Königsberg in latter-day Prussia was set on two sides of the Pregel River, and two islands between the banks were connected to each other and to the mainland by seven bridges. The problem was to identify a route through the city that would cross each bridge no more than once. Euler proved that there was in fact no solution to the problem. In doing this, he made a conceptual leap. This was in pointing out that the choice of route within each landmass is irrelevant, and that the only important element of the problem was in the sequence of bridges crossed. He thereby reformulated the problem abstractly, replacing land masses with abstract nodes, and bridges with abstract connections. The result was a mathematical structure called a *graph*; that is, a pictorial representation comprising a number of points (vertices) connected by arcs, or nonintersecting curves which can be distorted without affecting the graph. In carrying out the problem in this way Euler could deduce that, since the four land masses in the original problem are touched by an odd number of bridges, instances of a walk crossing each bridge only once results in a contradiction. If, on the other hand, Königsberg had one bridge fewer, thus an even number of bridges leading to a piece of land, a solution would have been possible.

Euler's totality of theorems and methods was huge. They include the demonstration of geometrical properties, for example, Euler's Line and Euler's Circle, new methods for solving quartic equations, the Prime Number Theorem, proofs of

Fermat's theorems and conjectures, the discovery of 60-plus amicable numbers, the calculus of variations and much more.

Development of Analysis, Number Theory, and Mechanics

Joseph Louis Lagrange (1736–1813) was a significant figure in Enlightenment mathematics, making contributions to the fields of analysis, number theory and mechanics. Lagrange was only nineteen when he solved the isoperimetrical problem—a problem which had been a subject of mathematical discussion for over half a century. He effected his solution by enunciating the principles of the calculus of variation and sent the answer off to Euler. Euler was so impressed that he withheld a paper he had been working on in the same area, so that Lagrange could complete his work and claim the invention of a new calculus. This placed Lagrange in the top rank of mathematicians then living. In 1758, Lagrange joined a society known as the *Turin Academy*, and it is in the volumes of its transactions that the bulk of his work can be found. In the first volume, he provides insights into the theory of the propagation of sound, indicating a mistake made by Newton. He obtained the general differential equation for the motion and integrated it for motion in a straight line. He also provides insights into recurring series, probabilities, and the calculus of variations. In the second volume, Lagrange sets forth a theory on the notation of the calculus of variations, illustrating its use by deducing the principle of least action, and by the solutions to several problems in dynamics. The third volume involves Lagrange's solution to dynamical problems using the calculus of variation along with papers on integral calculus. He also addresses the general differential equations of motion for bodies moving under mutual attraction.

Lagrange completed several works on algebra discussing representations of integers by quadratic forms and on algebraic forms in general. His papers of 1770 and 1771 put forward a general process for solving an algebraic equation of any degree via the Lagrange Resolvents. This method is significant for exhibiting the existing formulas for solving equations of second, third, and fourth degrees as manifestations of a single principle, a

method that formed the basis of Galois theory. In 1773, Lagrange proved the expression for the volume of a tetrahedron. Lagrange was also interested in number theory and was the first to prove that Pell's equation was incorrect.

It was for his seminal work on analytical mechanics, *Mecanique Analytique*, that Lagrange became best known. Written in Berlin in 1788, this offered the most comprehensive treatment of classical mechanics since Newton, forming the basis of mathematical physics in the 19th century. Lagrange based his research in mechanics on his own calculus of variations in which certain properties of a mechanistic system are inferred by considering the changes in a sum, or integral, that result from conceptually possible—or virtual—displacements from the path that describes their actual history of the system. From this, Lagrange identified independent coordinates necessary to the specification of a system comprising a finite number of particles, or *generalized coordinates*.

Lagrange later lectured at the École Polytechnique and in 1797 published papers on “Theory of Analytic Functions,” and “*Leçons sur le calcul des fonctions*,” or “Lessons on the Calculus of Functions.” These provided the basis of the first real textbooks on functions. In them, Lagrange attempted to substitute an algebraic foundation for the foundation of calculus, which was at that time problematic. Although he was ultimately unsuccessful his work spurred others on to develop the modern analytic foundation.

During his lifetime Lagrange was also heavily involved in the process of formulating new standard units of measurement at the close of the 18th century. It was Lagrange's influence that resulted in the final choice of the unit system of meter and kilogram and largely due to him that decimal subdivision was accepted in 1795 by the Commission for weights and measures in France.

Later Renaissance Mathematics and the Theoretical Explanation of the Solar System

Pierre Simon Laplace (1749–1827) was a French mathematician and astronomer best known for his investigation into the stability of the solar system. Laplace applied Newton's theory of gravitation to the solar system and managed to account for all observed deviations of the planets from

their theoretical orbits. He also demonstrated the usefulness of probability in the interpretation of scientific data.

It was in 1773 that Laplace began his life's work by undertaking to solve a particularly perplexing problem—that of Jupiter's orbit. Jupiter's orbit appeared to be shrinking continuously while that of Saturn expanded. The mutual gravitational interactions in the solar system were highly complex and seemed impossible to apply a mathematical solution to. For this reason, even Newton had concluded that divine intervention was required from time to time in order to preserve the system's equilibrium. Laplace set forth the invariability of planetary mean motions (average angular velocity) in 1773, and this was the first and most significant step since Newton toward establishing the stability of the solar system. Laplace examined the perturbations, or mutual gravitational effects, of planets and succeeded in proving that the eccentricities and inclinations of planetary orbits will always mutually remain small, self-correcting and constant.

Laplace applied quantitative methods to a comparison of living and inert systems. In 1780, along with Antoine-Laurent Lavoisier in 1780, and aided by a self-developed ice calorimeter, he illustrated that respiration was a form of combustion. In 1784 and 1785, Laplace worked on the subject of attraction between spheroids, work which resulted in his identification of the potential function of later physics. He examined the problem of attraction and discovered that the attractive force of a mass upon a particle, no matter the direction, can be obtained by differentiating a single function. This laid the mathematical foundation for the scientific study of magnetism, heat, and electricity.

In 1797, Laplace managed to remove the final remaining anomaly in the theoretical description of the solar system. He announced that lunar acceleration depends on the eccentricity of the earth's orbit. Whereas the mean motion of the moon depends on the gravitational attraction between the moon and Earth, this attraction is slightly diminished by the Sun's pull on the moon, solar action which depends itself on changes in the eccentricity of the earth's orbit, which is in turn affected by the perturbations of the other planets. The result is that the mean motion of the moon is accelerated so long as the earth's orbit becomes more circular. When the opposite happens, the

motion is retarded. This inequality is not truly cumulative, according to Laplace, but is of a period running into millions of years. Thus, Laplace removed the last threat of instability from the theoretical explanation of the solar system.

In 1796, Laplace published *The System of the World*, which was an exposition of his work in celestial mechanics. The book also included his nebular hypotheses. His *Celestial Mechanics*, which appeared between 1798 and 1827, summarized the results of his mathematical developments and application of the law of gravity. By calculating the motions of the planets, their satellites and perturbations, he devised a way of calculating the motions of the planets.

Laplace also published a popular work, *A Philosophical Essay on Probability*, which introduced his theory of probability to the masses. In it, he described the tools he had invented for mathematically predicting the probabilities that certain events will happen in nature. He applied his theory not just to ordinary events of chance but to the inquiry into the causes of phenomena, future events, and statistics. He also stressed its importance for astronomy and physics. This work also became renowned for its inclusion of Laplace's Central Limit Theorem in which Laplace proved that the distribution of errors in large samples of data taken from astronomical observations can be approximated by a normal distribution. A normal distribution is also known as a *Gaussian distribution* and is the most common distribution factor for randomly generated variables.

Laplace, like Lagrange, escaped imprisonment during the French Revolution, and was also involved in the organization of the metric system.

The Greenhouse Effect

Joseph Fourier (1768–1830) was a French mathematician and physicist renowned for illustrating how the conduction of heat in solid bodies could be analyzed in terms of infinite mathematical series. This became known as the *Fourier series*. Fourier's theory was the first accurate scientific theory to explain the diffusion of heat. Fourier was also known for his discovery of the greenhouse effect.

Fourier first developed an interest in mathematics at an early age and read Bezout's *Cours de*

Mathematiques at the age of just 14. In 1822, Fourier presented papers on the flow of heat in *Theorie analytique de la Chaleur* (*The Analytic Theory of Heat*). Based on Newton's law of cooling, Fourier deduced that the flow of heat between two adjacent molecules is in direct proportion to the minute difference in their temperatures. In his explanation, Fourier took three different aspects, one mathematical and the others physical. From a mathematical point of view, Fourier put forward the theory that any function of a variable, whether continuous or discontinuous, could be expanded in a series of sines of multiples of the variables. This was actually incorrect; however, his observation that some discontinuous functions equal the sum of infinite series was a revelation. Of the two physical elements of the theory, the proposal of his partial differential equation for conductive diffusion of heat is still taught to mathematical physicists today.

Fourier also began work on determinate equations, a work which was later completed by Claude-Louis Navier in 1831. In this, Fourier demonstrated his position on the roots of an algebraic equation. His work on the concept of the greenhouse effect rested on his calculation that the Earth, taking into consideration its size and distance from the sun, should be much colder than it is if warmed only by the effects of solar radiation. He suggested that a large portion of the additional heat received by the Earth was a result of interstellar radiation, a concept of the atmosphere that became known as the *greenhouse effect*.

Legacy of the Enlightenment Mathematicians

The mathematicians of the Enlightenment played a pivotal role in influencing many generations to come. They were responsible for the development of all main modern branches of mathematics as well as that of numerous other important disciplines in science. Isaac Newton, in addition to inventing the calculus, was credited with having invented the reflecting telescope and proposing a theory of light and color, developing three laws of motion, and devising the laws of universal gravitation. René Descartes was not only a creative mathematician but also an important scientific thinker. Descartes' mode of problem-solving and his

mathematical results were new and influential. There is also a philosophical significance to his work. By blending algebra and geometry and taking a novel approach to the geometrical status of curves, he made a significant contribution to ongoing philosophical debates surrounding then current mathematical practices.

Descartes can be credited with not only coframing the laws of refraction but also for developing an empirical account of a rainbow. More importantly, he offered us a new vision of the world, a world of matter possessing few material properties, and an immaterial mind, directly related to the brain. Descartes was a theorist who set the agenda for new directions in thought.

Joseph-Louis Lagrange made significant contributions to science and mathematics and transformed Newtonian mechanics into a branch of analysis. He also produced many important papers on problems in astronomy, and the mechanics of solids and fluids.

Pierre Simon Laplace was central to the development of mathematical astronomy and was also the first person to propose the existence of black holes. In addition, he was a central force behind systemizing probability theory, a theory that was to lay the groundwork for Bayesian statistics, and he pioneered the study of the speed of sound.

Joseph Fourier, an inspired mathematician, was a close friend of Napoleon Bonaparte and accompanied him on his Egyptian expeditions. Fourier helped to collate the scientific and literary discoveries they made during their time in Egypt. He was also later charged with constructing the new highway that ran from Grenoble to Turin, and it was at Grenoble that his mind turned to the important work he was to do on the propagation of heat.

Thora L. Fitzpatrick

See also Mathematics, 19th century; Mathematics, 20th century; Mathematics, Renaissance

Further Readings

Boyer, C. B. (1994). Analytic geometry: The discovery of Fermat and Descartes. In F. J. Swetz (Ed.), *From five fingers to infinity: A journey through the history of mathematics*. Open Court Publishing Company.

- Cooke, R. (1997). *The history of mathematics: A brief course*. Wiley.
- Katz, V. J. (2013). *A history of mathematics*. Pearson. Mathematicians of the seventeenth and eighteenth centuries. (n.d.). School of Mathematics, Trinity College, Dublin. <http://www.maths.tcd.ie/pub/HistMath/People/RBallHist.html>
- Motz, L., & Hane, J. W. (1993). *The story of mathematics*. Plenum Press.
- Rouse Ball, W. W. (2006). *A short account of the history of mathematics*. Scholarly Publishing Office, University of Michigan. (Reprint of 4th edition, original work published 1908)
- Struik, Dirk J. (1987). *A concise history of mathematics*. Dover Publications.

MATHEMATICS, MIDDLE AGES

Between the 5th and 15th centuries C.E., many significant scholarly developments occurred in the field of mathematics. This period has come to be termed the *Middle Ages* by contemporary historical researchers, and the intellectual achievements during this era, also referred to as the *medieval period*, have provided critical foundations for our modern mathematical and scientific understandings of the world. Four different cultural systems in particular experienced substantial mathematical progress during this millennium: Arabic culture, Chinese culture, European culture, and Indian culture.

These four cultures advanced their understandings of mathematics in contrasting ways. Some, such as the mathematicians of the Chinese and Indian cultures, developed their mathematical systems largely independent of outside influences, while others developed their systems based on the intellectual foundations provided by other cultures, in the way that European mathematicians adopted many elements from Arabic mathematics. Furthermore, these cultures progressed at different rates in their technical comprehension of the elements of mathematics. Chinese advancements in mathematics peaked during the 13th century, while the so-called Golden Age of Indian mathematics is said to have spanned between the 5th and 12th centuries. Likewise, Arabic mathematics in the Middle Ages climaxed between the 9th and 13th

centuries, whereas European mathematics peaked later, between the 12th and 15th centuries.

Examining the mathematical accomplishments of the Middle Ages can provide contemporary researchers with conceptual genealogies that will help to illuminate how we, in today's global civilization, have arrived at our modern mathematical theories of the universe. Although some concepts of the mathematicians of the Middle Ages may seem rudimentary or strange by modern standards, they have served as theoretical stepping-stones that scholars worldwide have used to help further our mathematical understanding. More specifically, the mathematicians of the Middle Ages made significant and novel advancements in the fields of algebra, geometry, and trigonometry, and many of their ideas are still taught today in classrooms worldwide.

Significant Arabic Mathematicians of the Middle Ages

Muhammad ibn Musa Al-Khwarizmi (780–850 C.E.) was a prolific Persian scholar whose range of knowledge extended over many fields of study, including algebra, astronomy, geometry, and translation. In a text he called *On the Calculation With Hindu Numerals*, Al-Khwarizmi translated into Arabic some of the theoretical works of the Indian mathematician Brahmagupta, and thus was responsible for being the first to introduce the Hindu numeration system and its concept of decimal places to the Arabic-speaking world, which European thinkers later translated from Arabic into Latin.

At some point during the 9th century, Al-Khwarizmi became the director of the prestigious academy known as the *House of Wisdom*, which was located in what is modern-day Baghdad, Iraq. It was during his time that Al-Khwarizmi developed and completed his most significant mathematical work, *The Compendious Book on Calculation by Completion and Balancing*. In this text, Al-Khwarizmi initiated the field of algebra by detailing the processes of first- and second-degree equations and how they could be practically applied.

Born in Baghdad, where he spent most of his life, Abu Bakr Muhammad ibn Hassan Al-Karaji (953–1029 C.E.) was another gifted Persian mathematician and engineer. Al-Karaji was interested

in applying his understanding of Arabic and Hindu mathematics to practical matters. For example, he wrote a technical manual called the *Book of the Extraction of Hidden Waters*, which made use of mathematics to describe the proper ways in which to dig wells and create canals.

Al-Karaji's most prominent work is his extensively studied *Al-Fakhri* (*The Glorious*, a book of algebra), in which he details arithmetic operations regarding polynomials, such as the procedures involving their addition, multiplication, and subtraction. In his other various mathematical works, Al-Karaji proved the binomial theorem, developed an algebraic usage of exponents, and attempted to differentiate algebra from geometry. Another important aspect of Al-Karaji's mathematical legacy is that he was the first to use the process of mathematical induction as a method of providing proof.

Hailed as one of the most preeminent scholars of the past two millennia, the Persian Abu Rayhan Al-Biruni (973–1048 C.E.) was expertly versed in a wide variety of fields, including astronomy, geology, history, linguistics, mathematics, physics, and theology. In the course of his long lifetime, Al-Biruni wrote prodigiously on the field of mathematics, publishing over 90 texts regarding mathematical subjects.

Perhaps Al-Biruni's most significant intellectual contribution to the field of mathematics is his work regarding mathematical geography as detailed in his masterwork, the *Determination of the Coordinates of Places for the Correction of Distances Between Cities*. In this text, Al-Biruni provides a process for which to determine latitudinal and longitudinal coordinates. To offer a practical application of this process, Al-Biruni utilized trigonometry to ascertain the direction of Mecca from the city in which he was then staying, Ghazna. Such an application allowed Al-Biruni to defend the importance of mathematics against the theological critiques of religious scholars of his era.

Significant Chinese Mathematicians of the Middle Ages

Yang Hui (1238–1298 C.E.) was a Chinese mathematician who hailed from the city that in modern times is known as *Hangzhou*. During his lifetime, Yang Hui wrote several texts on the topic of

mathematics, such as the widely studied works *Detailed Analysis of the Mathematical Rules in the Nine Chapters* and his *Fundamental Mutual Changes in Multiplication and Divisions*. However, Yang Hui is perhaps most widely renowned for his presentation of the concept known as *Yang Hui's triangle*, which provides a blueprint for binomial expansion calculations.

In fact, Yang Hui's triangle describes the same exact mathematical concept as Pascal's triangle, which Blaise Pascal organized four centuries later. As detailed in his various mathematical treatises, Yang Hui also made many other novel contributions to the field of Chinese mathematics, including processes for solving quadratic equations and discovering cube and square roots.

Born in the city of Yan-Shan, the Chinese mathematician Chu Shih-chieh (1249–1314 C.E.) is hailed by Eastern and Western scholars alike as one of the greatest mathematical minds to have ever lived. Such a statement becomes even more impressive when contrasted with the fact that Chu Shih-chieh is known to have written only two mathematical texts in his lifetime, his *Introduction to Mathematical Studies* and the *True Reflections on the Four Unknowns*.

For nearly 30 years, Chu Shih-chieh traveled copiously throughout both the northern and southern regions of China, and during his travels, he taught his mathematical ideas to countless students. Chu Shih-chieh made many significant developments to the field of mathematics, such as his method for calculating equations that involved decimals and fractions. He also developed new procedures for finding areas and volumes, solving polynomial equations, and calculating concurrent equations. In the centuries following his death, his two mathematical works became extremely popular theoretical textbooks in the territories of China, Japan, and Korea.

Significant European Mathematicians of the Middle Ages

Born in Rome to a family with a prestigious ancient lineage, Boethius (480–524 C.E.) was a highly educated Roman scholar. In fact, Boethius was so unusually well-educated for his time that contemporary researchers suspect that he likely studied in the famous scholastic cities of

Alexandria and Athens during his youth. Boethius's extensive education allowed him to translate the mathematical works of Euclid and Nicomachus into Latin, and his translations of these works served as the main mathematical textbooks available to Europeans until the 12th century.

Boethius's main mathematical text, *Arithmetic*, was based on Nicomachean modular arithmetic and served to introduce European audiences to Pythagoras's theories regarding natural numbers. Boethius's *Arithmetic* provided a large portion of the study materials which were utilized in European monasteries' educational courses of the era, the *quadrivium*, so-called, which was focused upon the fields of arithmetic, astronomy, geometry, and music theory. For this reason, Boethius was a critical forerunner of European mathematics.

Adelard of Bath (1080–1152 C.E.) was a prolific English scholar who is credited with being the first to introduce particular elements of Arabic, Greek, and Hindu mathematics to European thinkers. Adelard traveled widely in Europe and Central Asia during his lifetime, and during his travels, he mastered the Arabic, Greek, and Latin languages. His extensive knowledge of languages allowed him to translate important Arabic mathematical texts, such as the algebraic works of Al-Khwarizmi, into the European *lingua franca* of the period, Latin.

Furthermore, with certain elements of Arabic mathematics having been influenced by Hindu systems, Adelard is recognized for introducing the Hindu numeral system to Europe. Beyond these feats, Adelard wrote treatises regarding the proper usage of abacuses and astrolabes. Upon returning to Europe after his studies in Central Asia, Adelard settled in Laon, northern France, and began to give local lectures on the topics of Arabic mathematics and Arabic philosophy.

The Italian researcher Gherard of Cremona (1114–1187 C.E.) was an industrious translator of Arabic scholastic texts, many of which dealt with a variety of mathematical subjects. After becoming dissatisfied with the education he had received in Italy, Gherard moved to Toledo, Spain, with the precise intention of learning the Arabic language in order to read and study Arabic academic texts.

After mastering the Arabic language in Toledo, Gherard set out to translate as many Arabic scholastic books into Latin as he possibly could; consequently, he translated more than 80 Arabic texts

into Latin, of which the most significant were works concerning geometry and mathematics, such as the Arabic mathematician Ahmed ibn Yusuf's research regarding ratios. For his prodigious work in Arabic translation, Gherard is considered by modern historians as having been one of the fundamental conduits of Arabic mathematical concepts into European culture.

Leonardo Pisano Fibonacci (1170–1250 C.E.) in contemporary times is regarded as being perhaps the most brilliant of the European mathematicians of the Middle Ages. Born in Pisa, Italy, and educated in the Algerian region of North Africa, Fibonacci quickly realized that the Arabic–Hindu numeration system was vastly more efficient than the Roman system he had previously been exposed to. After completing his travels in North Africa, Fibonacci returned to Pisa and began work on a series of important mathematical texts.

The first of these works was Fibonacci's well-known *Book of Calculations*, in which Fibonacci expounds upon Arabic algebra and the Arabic–Hindu decimal place system. In this text, Fibonacci also details the famous *Fibonacci sequence*, which is a sequence of numbers in which each number in the series is the sum of the two numbers that precede it. Another of Fibonacci's famous mathematical works is his *Practice of Geometry*, wherein he presents a large assortment of geometrical problems and then proceeds to solve them using theorems derived from the ideas of the ancient Greek master, Euclid. These two works proved critical in the development of European mathematical understanding during the Middle Ages.

Born near modern-day Königsberg, Germany, the German polymath Regiomontanus (1436–1476 C.E.) was an expert in the fields of astronomy, mathematics, and translation. After coming into contact with a fragment of Diophantus's famous treatise, *Arithmetica*, Regiomontanus began a lifelong foray into the study of mathematics. As a result of his intense interest in astronomical topics, Regiomontanus began to develop trigonometrical procedures to aid his studies on astronomy.

In his work, *On Triangles of Every Kind*, Regiomontanus systematically develops a system of trigonometry. In this text, Regiomontanus describes and details concepts such as arcs, chords, circles, equality, ratio, the sine function, and quantity, and wherein he also declares more than 50

theories regarding geometry. Regiomontanus's work on trigonometry was therefore critical in providing European audiences with novel methods for solving triangles related problems.

Significant Indian Mathematicians of the Middle Ages

Hailing from Bhinmal, India, Brahmagupta (598–670 C.E.) was an astronomer and mathematician who became one of the most important Indian mathematicians of the Middle Ages. At some point in his adulthood, Brahmagupta became the director of the Ujjain astronomical observatory, which at the time was India's academic and mathematical epicenter. While at Ujjain, Brahmagupta wrote some of his most significant mathematical texts, such as his *The Opening of the Universe* and his *An Astronomical Treatise*.

In these works, Brahmagupta develops theories regarding algebra, the computation of square roots, latitudes, longitudes, and quadratic equations, but he is perhaps most famous for his novel ideas concerning the concept of zero. Brahmagupta described how the number zero is created when a number is subtracted from itself. He also detailed what occurs when zero is added to, subtracted from, or multiplied by a number, and these notions formed arithmetic methods that are still in use in contemporary times.

Bhāskara II (1114–1185 C.E.) was an Indian astronomer and mathematician who is reported to have lived during the peak of Indian mathematics in the 12th century. He became the director at the same Ujjain astronomical observatory that his intellectual forerunner Brahmagupta had been a leader of several centuries earlier. In his lifetime, Bhāskara wrote three important mathematical texts: *The Beautiful*; *Seed Counting or Root Extractions*; and his *Head Jewel of Accuracy*.

Within these texts, Bhāskara covered numerous mathematical concepts, wherein he displayed a mastery over the usage of the number zero as well as negative numbers. In his mathematical studies, Bhāskara realized that there were logical inconsistencies in Brahmagupta's conception of dividing numbers by zero, and thus, he revised this notion in a way that is reminiscent of modern methods. Furthermore, he developed novel methods for

squaring numbers and wrote extensive definitions on arithmetic, astronomical, and geometric terms.

William M. Peaster

See also Mathematics, Ancient; Mathematics, Renaissance; Medicine, Middle Ages; Physics, Middle Ages

Further Readings

- Abattouy, M. Muhammed Al-Karaji: A mathematician engineer from the early 11th century. *Muslim Heritage*. Retrieved January 2022 from <https://muslimheritage.com/muhammad-al-karaji-mathematician-engineer/>
- Adelard of Bath (2006), *Conversations with his Nephew: On the Same and the Different, Questions on Natural Science, and On Birds* Translated and edited by Charles Burnett with the collaboration of Baudouin van den Abeele, Pedro Mantas España & Italo Ronca. (Cambridge Medieval Classics, Series Number 9). Cambridge University Press.
- Anicius Manlius Severinus Boethius. School of Mathematics and Statistics at the University of St. Andrews. <https://mathshistory.st-andrews.ac.uk/Biographies/Boethius/>
- Conger, H., Overbay, S., & Schorer, J. Al Khwarizmi. The University of Kentucky. Retrieved December 2021 from <http://www.ms.uky.edu/~carl/ma330/project2/al-khwa21.html>
- Dallal, A. (2010). *Islam, science, and the challenge of history*. Yale University Press.
- Duhem, P. (1985). *Medieval cosmology, theories of infinity, place, time, void, and the plurality of worlds*. University of Chicago Press.
- Fibonacci, L. P. (1987). *The book of squares*. Academic Press.
- Gherard of Cremona. School of Mathematics and Statistics at the University of St. Andrews. <https://mathshistory.st-andrews.ac.uk/Biographies/Gherard/>
- Yang Hui. School of Mathematics and Statistics at the University of St. Andrews. <https://mathshistory.st-andrews.ac.uk/Biographies/Gherard/>
- Searcy, M. B. (1993, April). An exploration of Brahmagupta's formula using the geometer's sketchpad. *The Mathematics Educator*.
- Shivakumar, N. N., & Sridhar, S. (2014, March). A note on leading mathematician Bhaskara II of 12th Century. *International Journal of Innovative Technology and Research*.

Srinivasan, N. K. (2009, January). Madhava of Sangamagrama: The founder of mathematical analysis. *Scribd*. <http://www.scribd.com/doc/11509608/Madhava-The-Founder-of-Math-Analysis-Calculus>

Zhu Shijie. School of Mathematics and Statistics at the University of St. Andrews. <https://mathshistory.st-andrews.ac.uk/Biographies/Gherard/>

MATHEMATICS, 19TH CENTURY

During the 19th century, there was an unprecedented increase in the range and complexity of mathematical concepts. France and Germany, having been swept up in the tide of revolution that took over much of Europe during the latter part of the 18th century, produced some outstanding mathematicians. The two countries, however, took quite different approaches to the discipline. Post-revolutionary France under Napoleon put much focus on the practicality of mathematics. His reforms gave French mathematics a significant boost, as had been illustrated in the rise of the influential *three Ls* during the latter part of the 18th century: Joseph Louis Lagrange (1736–1813), Pierre-Simon Laplace (1749–1827), and Adrien-Marie Legendre (1752–1833).

The beginning of the 19th century saw Joseph Fourier's work on infinite sums reach fruition too. Fourier proposed the use of trigonometric functions as a tool for mathematical analysis. Fourier's series of periodic functions expressed as the sum of an infinite series of *sines* and *cosines* are still in use today in pure and applied mathematics. Fourier also helped to define what is meant by a function, although the definition in use today—that of defining it in terms of the corresponding elements of a domain and a range—is largely credited to German mathematician Peter Dirichlet.

This entry continues with discussions of the achievements of the most prominent figures in 19th-century mathematics—Évariste Galois, Carl Friedrich Gauss, Bernhard Riemann, George Boole, Georg Cantor, and Henri Poincaré—and, more briefly, the contributions of several other notable mathematicians of the period, including Charles Babbage, János Bolyai, Richard Dedekind,

Gottlob Frege, Nikolai Lobachevsky, Hermann Minkowski, August Ferdinand Möbius, Nicolò Paganini, George Peacock, and John Venn.

Évariste Galois

A towering figure in the world of 19th-century mathematics was Évariste Galois, who completed work in the late 1820s proving that there is no general algebraic method for solving polynomial equations of any degree greater than four. Galois's theory built on work previously been put forward a few years earlier by Norwegian mathematician Niels Henrik Abel. He also identified the impossibility of solving *quintic* equations, thereby solving a puzzle that had perplexed mathematicians for centuries. Galois paved the way for numerous developments including the origination of the field of abstract algebra, including algebraic geometry, rings, fields, group theory, modules, vector spaces, and noncommutative algebra.

Galois died in a duel at the age of 20, but in his short life, he had already made a name for himself in mathematical circles. At the age of 17, he began making important discoveries in the theory of polynomial equations (equations built from variables and constraints that used just the operations of addition, subtraction, multiplication, and non-negative whole number exponents). He succeeded, as noted earlier, in proving that it is not possible to have a formula for solving *quintic* equations. Quintic equations are polynomial equations with the term x to the power 5. Taking the work of Abel further, he also managed to prove that even more significantly, there are no algebraic methods in general to solve polynomial equations of any degree exceeding four. Galois derived these conclusions by examining whether or not the *permutation group* of its roots demonstrated a certain structure. Galois was the first mathematician to use the term *group* of permutations, thus setting the scene for the modern field of group theory. Galois's proof led to further definitive proofs or *disproofs* of the three classical problems set out by the ancient Greek philosopher Plato. This involved Galois in proving the impossibility of doubling of a cube and trisection of an angle. He also proved impossible the squaring of the circle.

Carl Friedrich Gauss

In Germany, by contrast, the influence of the great educationalist Wilhelm von Humboldt gave mathematics a rather different approach. Rather than simply supporting mathematics from a practical point of view, he encouraged the pursuit of pure mathematics for its own sake.

It was in this environment that Carl Friedrich Gauss came to prominence, a mathematician so esteemed that he was sometimes referred to as the *Prince of Mathematics*. In many ways, Gauss's ideas were ahead of their time, touching as they did on a wide variety of mathematical disciplines including geometry, number theory, algebra, calculus, and probability. One of the greatest mathematicians of all time, he is put in the same category as Archimedes and Isaac Newton.

Gauss was a veritable child prodigy, having already mastered and begun to criticize Euclid's geometry by the age of 12. His intellectual prowess attracted the attention of a wealthy benefactor, the Duke of Brunswick, who paid for him to attend the University of Göttingen, and it was here that he discovered several important theorems. Gauss was the first mathematician to identify any pattern in the occurrence of the prime numbers. This was a problem that had plagued mathematics since ancient times. Even though the occurrence of prime numbers appeared to be completely random, Gauss took an innovative approach to the problem by graphing the incidences of primes as the numbers increased. He noticed a trend or rough pattern whereby as the numbers increased by 10, the probability of prime numbers occurring reduced by a factor of roughly 2 (e.g., there is a one in four chance of there being a prime in the numbers one to 100, a one in six possibilities of getting a prime in the numbers one to 1,000, and a one in eight chances of a prime in numbers one to 10,000). Gauss was perfectly aware, however, that this method was only an approximation, and for this reason, he kept his early findings secret until he was far older.

At the age of 19, Gauss constructed what had previously been considered impossible—a regular 17-sided polygon, by using just a ruler and compass. This proved that a heptadecagon could be constructed geometrically, and this was a major mathematical advance even since the time of

Greek mathematics. In the same year, he formulated his prime number theorem on the distribution of prime numbers among the integers, proving that every positive integer can be represented as a sum of at most three triangular numbers.

Gauss proposed the first clear explanation of complex numbers and a clear exposition of the functions of complex variables early in the 19th century. Imaginary numbers, involving i , the imaginary unit equal to the square root of -1 , had been in use since the 16th century to solve previously unsolvable equations. Leonhard Euler (1707–1783) later completed ground-breaking work on these and on complex numbers but failed to provide a clear picture of how imaginary numbers could be connected to real numbers. Jean-Robert Argand (1768–1822) was one of the first mathematicians to use graphs to depict and interpret complex numbers, but Gauss was later responsible for popularizing this practice. He also went on to introduce the standard notation $a + bi$ for complex numbers. This subsequently resulted in the significant expansion of the theory of complex numbers. Gauss was also soon to prove what is now called *The Fundamental Theorem of Algebra*, which was not in fact directly related to algebra at all. The theorem illustrated that the theory of complex numbers is an algebraically *closed* system. This is in contrast to real numbers, and it holds that the solution to polynomials that have real coefficients can yield solutions in the complex number field.

In 1801, Gauss published *Disquisitiones Arithmeticae*, one of the most influential mathematical works ever written. This provided the foundation for modern number theory, provided an exposition of modular arithmetic, and presented the first proof of Euler's and Legendre's law of quadratic reciprocity.

Gauss retained an immense interest in theoretical astronomy for much of his life and was appointed the director of Göttingen's astronomical observatory, a post in which he served for many years. He predicted the position of the planetoid Ceres, a prediction that was quite different from those of his contemporaries, and he was later to be proven correct. To arrive at this prediction, Gauss had used one of the first applications of the least-squares approximation method.

Gauss also introduced what is now termed the *Gaussian distribution* into the areas of probability and statistics. This, along with the Gaussian function and the Gaussian error curve, illustrated how probability could be shown as a bell-shaped or *normal* curve, which peaks at the mean or expected value and then falls off quickly toward plus or minus infinity. This probability curve is fundamental to descriptions of statistically distributed data. He also performed the first systematic study of modular arithmetic drawing on modulus and integer division. Today, this is used in abstract algebra, number theory, computer science, and the visual and musical arts.

Another inquiry among the many that were pursued by Gauss was the shape of the Earth. Gauss had begun to speculate even on the shape of space itself, leading him to question one of the most central theories in the whole of mathematics, Euclidean geometry. The premises of Euclidean geometry lie on a flat, not curved, universe. Gauss, probably quite sensibly, decided not to publish his thoughts in this area, however, having seen what had happened to others, most notably Galileo, when they set forth avant-garde notions. He left the field open to János Bolyai and Nikolai Lobachevsky, who were willing to enter this controversial area.

Gauss was highly interested in the field of differential geometry, which is the field of mathematics that deals with curves and surfaces. He developed what became known as *Gaussian curvature*, which provides an intrinsic measure of curvature, which depends solely on how we measure distances on the surface rather than how they are embedded spatially.

Gauss was not only a consummate mathematician; he also invented the heliotrope, an instrument for marking positions in land surveys, and he worked with Wilhelm Weber on the measurement of the Earth's magnetic field. His work in electromagnetism earned him recognition in terms of the international unit of magnetic induction being named the *gauss*.

As his fame spread throughout the world, Gauss was the mathematician of choice when it came to complex mathematical questions. In his later years, he was criticized for dismissing the ideas of younger talented mathematicians and was accused of attempting to claim their ideas as his own.

Bernhard Riemann

The German mathematician Bernhard Riemann, like Gauss, was working on a type of non-Euclidean geometry termed *elliptic geometry*. He was also involved in work on a generalized theory of geometry. During his work, Riemann broke away from two- and three-dimensional geometry as being too limiting and began instead to think in terms of higher dimensions.

Riemann attended Gauss's lectures and was one of the lucky few to actually benefit from his patronage. Riemann's form of non-Euclidean geometry supposed that there are no such things as parallel lines and that the angles of a triangle do not add up to 180° . His advances in geometry and his novel concepts unified and generalized the various existing forms of geometry. He also generalized the concept of mathematical space, curves, and surfaces. At the age of 26, Riemann gave a lecture on the foundations of geometry, outlining his version of mathematics, which supposed that there are many different kinds of space. Hitherto, we had envisioned ourselves as inhabiting just one type of space, the flat Euclidean type of space. Riemann also put forth his one-dimensional complex manifolds, termed *Riemann surfaces*. These theories fundamentally changed how we see the world and paved the way for the development of higher dimensional geometry.

Riemann's metric left the limitations of two- and three-dimensional geometry behind, and even Bolyai and Lobachevsky's curved spaces, in favor of higher dimensional geometry. He extended the differential geometry of surfaces into n dimensions and devised a conception of a multidimensional space. This became known as *Riemannian space*, and it was this concept that allowed for the future development of the theory of relativity, which forms the basis of much of modern-day mathematics. Riemann was responsible for introducing the collection of numbers known as *tensors* into mathematics. These applied at every point in space and described to what extent space was curved or bent. For example, in four spatial dimensions, a group of 10 numbers is required at each point to describe the mathematical space's properties, no matter the level to which it is distorted.

Riemann is perhaps best known for his work on a function in the complex plane termed the

Riemann zeta function. This builds on the simpler zeta function created earlier by Euler. Riemann used this function to build a three-dimensional landscape and recognized that this theoretical landscape could be used to unlock the fundamental, longed-for solution to prime numbers. He noticed that the surface of his three-dimensional graph dipped to height zero at key places. This enabled him to demonstrate that the first 10 zeroes, for some unknown reason, appeared to line up in a straight line through the three-dimensional landscape of the zeta function. This became known as the *critical line* and the observation became known as the *Riemann hypothesis*. This theoretical leap took Riemann to the realization that these zeroes were connected in a completely unexpected manner to the way prime numbers are distributed. Riemann calculated that he could perhaps use this as a framework to correct Gauss's guesswork, in which he regarded the number of primes as numbers when counting higher and higher. Riemann's hypothesis, which remains unproven, purports that all the zeroes would theoretically be on a straight line and provides a strong indication that Gauss's approximations were accurate and that the primes are distributed over the universe of numbers in a balanced, beautiful, and regular way.

Riemann's hypothesis of the zeta function was published in 1859, but, unfortunately, many of his papers on the topic were destroyed in a fire after his early death, aged just 39. Today, the Riemann hypothesis is still one of mathematics' greatest unsolved questions, and a prize of over \$1 awaits the first person who can offer a final solution.

Boolean Algebra and Logic

In the mid-19th century, the British mathematician George Boole created a system of algebra, which is now known as *Boolean algebra* or *Boolean logic*. In his system, the only operators were *and* or *and not*, and these alone could be used to solve logical problems and mathematical functions. Boole also devised a binary system, which depended on just two objects: *on or off*; or *true or false*; or 0 or 1. In this system, Boolean algebra was famously born: $1 + 1 = 1$. Boolean algebra provided the basis for modern

mathematical logic, ultimately leading to the development of computer science.

Boole was one of the few mathematicians since Gottfried Wilhelm Leibniz (1646–1716) to seriously consider logic and its implications. Boole, unlike Leibniz, however, looked at logic as a discipline of mathematics rather than of philosophy. Unlike many other mathematicians at the time, Boole did not demonstrate his talent early in life. He was 34 before he published anything of note. He first began to see the possibilities of using algebra for solving logical problems by pointing to the deep analogy between the symbols of algebra and those of logic in terms of forms and syllogisms. He became determined to devise and develop a system of algebraic logic that would even go so far as to define and model the brain's function. He professed a profound confidence in logical reasoning, speculating a *calculus of reason* in the 1840s and 1850s.

Boole set out to find a way to encode logical arguments in a language that allowed them to be solved mathematically. His form of linguistic algebra, now termed Boolean algebra, involved three basic operations. These are *and*, *or*, and *not*; and Boole saw these as the only necessary operations required to perform comparisons of things and to carry out basic mathematical functions. The use of symbols and connectives enabled Boole to simplify logical expressions, for example, the algebraic identities of $(X \text{ or } Y) = (Y \text{ or } X)$, $\text{not}(\text{not } X) = X$, $\text{not}(X \text{ and } Y) = (\text{not } X) \text{ or } (\text{not } Y)$, and so on. He also devised a new approach based on a binary system, which involved processing just two objects: “yes–no,” “true–false,” “on–off,” and “zero–one.” If “true” is represented by 1 and “false” is represented by 0, then under the terms of Boolean algebra, it is possible for $1 + 1$ to be equal to 1 (where “+” is an alternative representation of “or”).

During Boole's lifetime, his ideas were largely ignored or criticized. It was not until the American philosopher, mathematician, and logician Charles Sanders Peirce (1839–1914) took them up and elaborated on them that they received the recognition they deserved, and this was several years after Boole's death in 1864. Then, more than 70 years later, Claude Shannon (1916–2001) made the connection between Boole's theories and their application to the processes and mechanisms of the real world. In fact, Boole's system is used nowadays in

the electromechanical relay circuits that underlie modern electronics and digital computers.

An Irish mathematician, William Rowan Hamilton (1805–1865), extended the concept of number and algebra further, in 1843 coming up with the theory of *quaternions*, a four-dimensional number system. This theory asserted that a quantity representing a three-dimensional rotation can be described by just an angle and a vector. Quaternions provided the first example of a noncommutative algebra, where $a \times b$ does not always equal $b \times a$, thus showing that several different consistent algebras may be derived by selecting alternative sets of axioms.

Hamilton's quaternions were later extended and developed by Arthur Cayley (1821–1895), who was also a pioneering mathematician in the area of modern group theory, higher dimensional geometry, matrix algebra, and the theory of invariants.

During the 19th century, mathematics became more abstract and complex; however, it also saw some mathematicians returning to their roots, with the emphasis on mathematical rigor. In the early part of the century, Bernard Bolzano (1781–1848) was one such mathematician, who put the focus on analysis and came up with the first pure analytical proof of the fundamental theorem of algebra and the intermediate value theorem. He also made early forays into the world of mathematical sets or collections of objects defined by a common property, for example, all numbers greater than 5.

Riemann and the French mathematician, engineer, and physicist Augustin-Louis Cauchy (1789–1857) rigorously reformulated calculus, which led to the development of mathematical analysis. This was a system of pure mathematics involved with the notion of limits. These could be limits to sequences or functions and with the theories of integration, infinite series, and differentiation. In 1845, Cauchy proved a theorem (known as *Cauchy's theorem*) that was a fundamental group theory. The German mathematician Carl Jacobi (1804–1851) made significant contributions to analysis too, as well as in the areas of determinants, matrices, and periodic functions.

Georg Cantor

In the latter part of the 19th century, Georg Cantor made a significant impression on mathematics.

The German mathematician's first papers were on number theory, but he was soon to turn his attention to calculus or analysis. He aimed to solve an ongoing problem on the uniqueness of the representation of a function by trigonometric series. What he remains best known for, however, is being the first mathematician to grasp the meaning of infinity and accord it a mathematical precision. Galileo had attempted to tackle the concept of infinity back in the 17th century along with the difficulties thrown up by comparatively different infinities. He managed to show that a one-to-one correspondence could be drawn between all the natural numbers and the squares of all the natural numbers to infinity. This suggested that there were as many square numbers as integers which goes against common sense since we intuitively know that there are many integers that are not squares. This concept became known as *Galileo's paradox*. Galileo also revealed that two concentric circles must both be made up from an infinite number of points, although the larger circle would seem to contain a larger number of points. Galileo eventually admitted defeat, concluding that concepts such as less, greater, and equal to could be applied only to finite sets of numbers rather than infinite sets per se. This was Cantor's challenge, and he first took it up by stating that if it were possible to add one and one, or 20 and 20, and so on, it ought to be possible to add infinity and infinity. He understood that it was possible to add and subtract infinities and that there existed another larger, then yet another larger infinity beyond the standard notion of infinity.

This idea of an *infinity of infinities* clearly had philosophical as well as mathematical implications, and the theory would change the approach to mathematics forever. Cantor's theory derives from his consideration of an infinite series of natural numbers (1, 2, 3, 4, etc.) and then an infinite series of multiples (15, 30, 45, etc.). He saw that while the multiples of 15 were a subset of the natural numbers, the two series could be paired, one-to-one. For example, 1 could be paired with 15, 2 with 30, and so on. This process is termed *bijection*, and it illustrates that both sets were in effect the same size because they had the same number of elements. This can clearly apply to other subsets of natural numbers, for example, the odd numbers 1, 3, 5, 7, and so on, or to squares 1,

4, 9, 16, and so on. It could even apply to sets of negative numbers and integers. Cantor found that he could even pair up all fractions or rational numbers with whole numbers, revealing that rational numbers possessed the same sort of infinity as natural numbers. This was out of line with the fact we intuitively feel there must be more fractions than whole numbers.

The difficulty lay in considering decimal numbers, which include irrational numbers such as π , e , and $\sqrt{2}$. He came up with some ingenious arguments, for example, the *diagonal argument*, to prove it was always possible to construct a new decimal number that was missing from the original list. This provided the evidence that the infinity of decimal numbers was bigger than the infinity of natural numbers. Cantor also managed to show that these numbers were *uncountable*, which meant they contained more elements than it would ever be possible to count. This is in opposition to rational numbers, which were technically *countable*. Cantor used the term *transfinite* to distinguish the levels of infinite numbers from an absolute infinity, which he equated with God. While the size of a finite set is simply a natural number stating the number of elements in the set, there needed to be a form of notation to describe the sizes of infinite sets.

Cantor's revolutionary theory on infinity heralded a new era in mathematics. It also led to the idea of the *continuum hypothesis*, whereby there can be other infinities, or even lots of infinities, between the infinity of whole numbers and the bigger infinity of decimal numbers. Cantor himself did not believe there was any such infinite set. The continuum hypothesis was, however, an ongoing puzzle for another century when it was taken up by Julia Robinson and Yuri Matiyasevich in the mid-20th century.

While Cantor's work on the concept of infinity was a landmark in mathematics, he also completed important work in set theory. This work has since become integral to modern-day mathematics throughout all the various branches that have grown up. The concept of a set had been implicit since the dawn of mathematics in Aristotle's day; however, it was limited to general finite sets. The infinite was not considered suitable for mathematics and was kept to a purely philosophical discussion. Cantor, however, succeeded in showing that there could in fact be infinite sets of

different sizes, just as there were finite sets. Some of these would be *countable* and some *uncountable*. During the latter part of the 19th century, Cantor worked on refining his set theory and produced definitions of *well-ordered sets* and *power sets*. This led to the introduction of his concept of ordinality and cardinality. What we now known as *Cantor's theorem* pronounces that

For any set A , the power set of A (set of all A 's subsets) has a cardinality which is greater than A .

This means the power set of a *countably* infinite set is uncountably infinite.

Although Cantor's set theory has become a part of today's mainstream mathematics, it was widely mistrusted by Cantor's peers. Some theologians saw his discussion of the infinite as an attack on their view of the nature of God. Others saw it as the willful destruction of the whole basis of mathematics. This included mathematician Henri Poincaré, who compared Cantor's set theory to a disease. There were some, however, who could see its potential, including David Hilbert, who declared Cantor's set theory to be a *paradise*.

As Cantor grew older, he succumbed to mental illness, which some viewed as a result of spending so much time contemplating such complex, abstract, and paradoxical concepts. He became obsessed with the notion that it was Francis Bacon who had written the Shakespearean plays and that Jesus was in fact the natural son of Joseph of Arimathea. He died alone in his room at the Halle sanatorium, leaving his life's project unfinished.

Henri Poincaré

Henri Poincaré was a leading light in the world of mathematics at the end of the 19th century and excelled in virtually all fields from geometry and algebra to analysis. He was a big believer in intuition and saw it as just as important as logic in the process of discovery. Poincaré is well known for his partial solution to a puzzle that had plagued the mathematicians Euler, Lagrange, and Laplace. This was termed the *three-body-problem*. Isaac Newton had shown that the paths of two planets in orbit around each other would stay stable, but the addition of even one more orbiting body would result in the involvement of up to 18 differing variables,

including position, velocity in either direction, and so on. This would make it mathematically too difficult to be able to predict or disprove a stable orbit.

Poincaré provided a solution to this vexing problem by setting out a series of approximations of the orbits. While this did not entirely solve the problem, it did enough to win him a generous prize from the King of Sweden in 1887.

In retrospect, Poincaré realized he actually had been mistaken. His simplifications did not imply a stable orbit, and in fact, just a slight deviation from his initial conditions would result in vastly different orbits. This realization, despite being born from error, resulted in a new theory now known as *chaos theory*, which is well known for being the theory which implies that the fluttering of a butterfly's wings could lead to a tornado on the other side of the planet. The theory was the first to indicate that there is a threshold for chaotic behavior. Poincaré's mistake, instead of diminishing his reputation, served to enhance it. He went on to develop the science of topology, first instigated by Leonhard Euler with his "Seven Bridges of Königsberg" problem. Topology is a form of geometry involving one-to-one correspondence of space, and it is also sometimes called *bendy geometry* since it implies that two shapes are the same if one can be morphed into the other without any cutting. Topology would state that a football and a banana are topologically equivalent. And so are a doughnut, which has a hole in the middle, and a teacup, with its handle. However, a doughnut and a football differ topologically because one cannot morph into the other. Poincaré produced a list of all possible two-dimensional topological surfaces. However, when it came to describing the three-dimensional universe, he came up with the *Poincaré conjecture*. This vitally important theory came to dominate mathematics for close to a century.

Poincaré also worked in theoretical physics, providing a presentation of the Lorentz transformations that formed an important step in the development of Einstein's theory of special relativity. He also contributed to areas such as optics, electricity, elasticity, quantum theory, and cosmology. Poincaré the engineer was also one of the last great mathematicians to adhere to a traditional view of mathematics, which felt that human intuition was as important as formalism and rigor. Often called the *last universalist*, he was one of the

last great mathematical all-rounders. The 20th century would see the rise of the specialist in mathematics, since the field had become so much more varied and complex.

Other Notable Mathematicians of the 19th Century

This entry has briefly outlined some of the mathematicians who defined the 19th century. There were many others who made significant contributions to mathematics during this time. One name immediately stands out, that of János Bolyai, an accomplished mathematician who had studied under Gauss. Bolyai was obsessed with Euclid's fifth postulate, which had been a standard geometrical principle for over 2,000 years. The principle stated that only one line can be drawn through a given point, so that the line lies parallel to a given line that does not contain the point. Additionally, its corollary is that the interior angles of a triangle amount to 180° or two right angles. Even through his obsession with the problem was starting to make him ill, Bolyai continued to pursue his quest, ultimately concluding that it was possible for there to be consistent geometries that did not require Euclid's parallel postulate. Bolyai's imaginary, "hyperbolic" geometry explained the geometry of curved spaces on saddle-shaped planes where the triangle's angles did not add up to 180° , and seemingly parallel lines were not in fact parallel.

Although we can visualize a flat surface and a surface with a positive curvature, for example, a sphere such as planet Earth, it is impossible to imagine a hyperbolic surface with a negative curvature, unless it is over a small area where it would resemble a saddle. The concept of a hyperbolic surface seemed to go against reality and was a radical departure from Euclid's geometry. Bolyai's treatise was published in 1832 and was obviously well regarded by Gauss who, having read it, tried to claim the idea as his own.

Independently of Bolyai in the Russian city of Kazan, Nikolai Ivanovich Lobachevsky also developed a theory of geometry which was non-Euclidean. Lobachevsky had also been working on other significant mathematical concepts including a method for approximating the roots of algebraic equations and the definition of a function in terms of the

correspondence of two sets of real numbers. Lobachevsky, like Bolyai, also died penniless, and in obscurity, Bolyai having succumbed to the mental illness he had earlier developed as a result of his mathematical obsession.

Meanwhile, in Britain, mathematics was undergoing somewhat of a resurgence. Earlier in the century, Charles Babbage, a mathematician, philosopher, inventor, and engineer, had designed a machine that could not only automatically perform computations based on a program of instructions stored on cards but also could compute logarithms and trigonometric functions. It was not until 100 years later that a working model of his machine was made and found to work perfectly. During the 19th century, the Englishman George Peacock is credited with inventing symbolic algebra, and he is also renowned for extending the scope of algebra beyond the ordinary number system. Recognition of the possible existence of nonarithmetical algebras was an important step toward future developments in abstract algebra.

The German mathematician and theoretical astronomer August Ferdinand Möbius (1790–1868) gave his name to several discoveries including the *Möbius strip*, the *Möbius configuration*, the *Möbius function*, and the *Möbius inversion formula*. Marius Sophus Lie was primarily responsible for the theory of continuous symmetry and its application to the geometric theory of differential equations. Nicolò Paganini discovered the second smallest pair of amicable numbers in 1866, something entirely overlooked by mathematicians even as great as Euler. Richard Dedekind defined concepts such as similar sets and infinite sets and devised a notion that is now known as the *Dedekind cut*, which showed that an irrational number divides the irrational numbers into two sets or classes. The upper set is strictly greater than all members of the lower set. This means that every location on the number line continuum contains either a rational or irrational number, and there are no gaps or empty locations. This theory is still used as a definition of real numbers. In 1881, John Venn introduced *Venn diagrams*, which remain one of the best-known and most-used tools in set theory.

In 1896, two proofs for the Prime Number Theorem were presented, by Jacques Hadamard and Charles de la Vallée Poussin. These proofs

were developed independently but both illustrated that the number of primes occurring up to any number x is asymptotic (tends toward) $\frac{x}{\log x}$.

Hermann Minkowski developed a system of number theory termed *geometry of numbers* at the end of the 19th century. This was a multidimensional geometrical method for solving number theory problems. Complex concepts such as lattice points, convex sets, and vector spaces formed part of the theory. In 1907, Minkowski discovered that Einstein's special theory of relativity was better explained in a four-dimensional space. This is often known as *Minkowski space-time*.

In 1879, the influential mathematician Gottlob Frege (1848–1925) produced the *Begriffsschrift* (*Concept-script*), which set the stage for a new field of logic, especially in its rigorous approach to the treatment of functions and variables. Frege's attempt to prove that mathematics is a product of logic involved him in devising techniques far beyond the traditional logic of Aristotle. He was the first to introduce the notion of variables into logical statements, along with *quantifiers*, *existentials*, and *universals*. Frege took Boole's propositional logic a stage further by extending it into *predicate logic*. His groundwork paved the way for the radical intellectual advances of Bertrand Russell and Giuseppe Peano.

Conclusion

The 19th century saw an explosion in applied mathematics as the industrial revolution gathered pace throughout Europe. The turbulent French Revolution (1789–1799) had been followed by the development of the École Polytechnique in Paris. There the finest mathematicians in the country developed their skills in mathematics. Many graduates from the polytechnic went on to excel in specialized forms of engineering in terms of the building of roads, mines, and creating artillery.

During the 19th century, many national mathematical societies became no more including the London Mathematical Society in 1865, the Société Mathématique de France in 1872, and the American Mathematical Society in 1888.

Thora L. Fitzpatrick

See also Geometry, Classical; Geometry, Non-Euclidean; Mathematics, 20th century; Mathematics, Antiquity;

Mathematics, Enlightenment; Mathematics, Middle Ages; Mathematics, Renaissance.

Further Readings

- BBC. (2021). A brief history of mathematics. *BBC Radio 4*. <http://www.bbc.co.uk/programmes/b00srz5b/episodes/guide>
- Bellos, A. (2010). *Here's looking at Euclid*. Free Press.
- Bonola, R. (1955). *Non-Euclidean geometry: A critical and historical study of its development* (English translation with additional appendices by H. S. Carslaw). Dover.
- Boyer, C. B. (1991). *A history of mathematics* (2nd ed.). Wiley.
- Grattan-Guinness, I. (2000). *The search for mathematical roots, 1870–1940: Logic, set theories and the foundations of mathematics from Cantor through Russell to Gödel*. Princeton University Press.
- Kolmogorov, A. N, & Yushkevich, A. A. (Eds.). (1996). *Mathematics of the 19th century*. Birkhäuser. <http://dx.doi.10.1007/978-3-0348-9173-8>
- Magnani, L. (2001). *Philosophy and geometry: Theoretical and historical issues*. Kluwer.
- Yaglom, I. M. (2009). *Geometries, groups and algebras in the nineteenth century—A history*. Ishi Press.

MATHEMATICS, RENAISSANCE

The triumph of the postulational form of thinking based on deductive reasoning and its symbolic representations had been the characteristic features of mathematics during the Renaissance—a period of incredible cultural and intellectual advancement based on the revival of classical learning, which spread all over Europe in the 14th, 15th, and 16th centuries. Mathematics during this time had two distinct trends—the rise in arithmetical–algebraic studies and the development of the theory of fluxions (flowing or changing quantity, as it was termed by Isaac Newton). The mode of expression had changed from a mere set of rules for field workers or students to postulations for analysts with proper deductions and justifications (composing proofs).

The invention of the printing press around 1450 made it possible for hitherto unfamiliar classical works to be available to the scholars of 15th-century Italy, the cradle of the Renaissance, more

than ever before. The initial thrust of the Latin translations of Arabic algebraic and arithmetical treatises obscured the Greek mathematical works for a long time. Explorations of the ancient Greek works at that time were difficult because few people read classical Greek, and the translated works required advanced training in mathematics. Later, as 17th-century mathematicians resorted more and more to the Greek literatures, after overcoming their initial handicaps, mathematical thinking attained new heights through the ingenious use of extensive and elaborate expositions of Greek arithmetic and geometry, thus far untried by Western scholars. The word *mathematics* itself has its origins in the Greek word *manthanein*, used by Pythagoras (c. 569–c. 475 B.C.E.) to denote someone with exceptional and wide learning. The word also carries a meaning of being alert, awake, wise, and inquisitive. Renaissance mathematics, with all its emphasis on logic and specificity, evolved not only out of the intellectual achievements of ancient Greek civilization but also from the contributions of Mesopotamian, Chinese, Indian, and Islamic civilizations. Historian of science Peter Dear proposed a two-phase model for the understanding of the evolution of early modern science. The first phase, the scientific renaissance of the 15th and 16th centuries, was a period of restoration of the knowledge of nature. This phase gradually paved the way for the second phase—the Scientific Revolution of the 17th century—when scholars shifted their attention from recovery of classical learning to the innovations of the modern paradigms.

The Renaissance and Greek Mathematics

The ancient Greek scholars were the pioneers in the scientific analysis of applied knowledge. They added the element of logical structure to the study of geometry, raising it above a mere practical science needed for such mundane things as land acquisition and distribution. The Greeks were distinctive for their individual excellences.

Thales (c. 624–c. 546 B.C.E.), for example, is often considered as the first mathematician to whom several individual results are ascribed. He is known for his deductive methodology. Thales is also remembered for the first truly mathematical demonstration of the postulations, pioneering the

general way of understanding the geometrical (mathematical) truths that became the foundation of Renaissance mathematics. Thales's approach was more logical and comprehensive than the existing practice of formulating mere algorithmic statements or instructions. Pythagoras (c. 569–c. 475 B.C.E.) was rather a mystic and aristocratic figure with a scientific mind. The terms *philosophy* (“love of wisdom”) and *mathematics* (“that which is learned”) are known to have been coined by him. Pythagoras and his secret brotherhood demonstrated several theorems and constructions of geometry. The extant works on conic sections by Apollonius (c. 262–c. 190 B.C.E.) prepared the way for the astronomical inventions of the Scientific Revolution almost two millennia later. A summation of the years of Greek practices, thoughts, and improvements in mathematical knowledge can be seen in the *Elements* of Euclid (c. 300 B.C.E.). All the harmonies between the theories and realizations of arithmetic and geometry were upset, however by some problems posed by Zeno (c. 490–c. 430 B.C.E.) involving the infinite and the infinitesimal. In attempts to set right the instabilities that Zeno had rendered, Archimedes (c. 287–c. 212 B.C.E.) held up the exhaustion method that avoided the pitfalls of infinitesimal by simply discarding them. He was exceptionally bold among the Greeks in handling questions of infinity and the infinitesimal. The sophistication attained by the ancient Greeks was far beyond the reach of later generations of European mathematicians until the 17th century. The Greek scholars, however, made few significant advances except in geometry and trigonometry. This shortfall is sometimes ascribed to intellectual obstacles between Greek geometry, with its focus on whole numbers, and the roots of algebra. From its origins in ancient Egypt and Babylonia, however, algebra continued to develop in Asia, and this allowed Indian and Arabic scholars to act as an intellectual bridge between medieval and modern mathematical studies, riding upon the political wave of the Islamic invasion of the West.

Influences of the Islamic Tradition

The most significant historical development affecting Renaissance mathematics was the spread of Islam. Both Muslim and non-Muslim medieval scholars, living under Islamic rule, played a crucial

role by transmitting the Greek, Hindu, Chinese, and other pre-Islamic fruits of knowledge to Western world. It all started under the caliphate of al-Mamun (786–833), who set up the “House of Wisdom” in Baghdad. The initial stress was on the translations of all the important literatures collected from Persia, Greece, China, and India through trade connections and political treaties with the Roman Empire. Soon the Arabian scholars started analyzing and interpreting those texts with profound insights that opened new horizons in the history of mathematical thought.

Mohammed ibn Musa al-Khwarizmi (780–850), a noted scholar of Baghdad, started translating the Indian texts. His first book was an exposition of the Hindu scheme of numeration, which later enormously influenced Renaissance mathematicians. This scheme later came to be known after him as *algorism* or *algorithm*. Any given rule or procedure of operation is now called an algorithm. Perhaps, the most important work of al-Khwarizmi was *Al-kitab al-mukhtasar fi hisab al-jabr wa'l muqabala* (*The Compendious Book on Calculation by Completion and Balancing*), which gave a straightforward exposition of the solution of equations, especially of the second degree. *Al-jabr* contained rules for the operations on binomial expressions and the rules governing the negative numbers, though Khwarizmi rejected the negative roots as solutions. The deductive way of thinking and the structurally comprehensive mode of expressions with a distinct beginning and a precise conclusion were the features of the writings of the Arabian scholars that would later influence the theoretical structures of Renaissance mathematics. The rhetoric form of expression of the Arabian scholars, however, did not attract the Renaissance scholars of the West. The Arabic mathematical works were ramified in four major directions, all of which were ultimately formalized into a logically connected abstract mathematical schema emphasizing the postulational form of thinking, deductive reasoning, and symbolic expression in the following centuries. These were:

Numeration—Based on the Indian principle of positions and Greek or Roman form of numerals.

Algebra—Derived from the neighboring civilizations of Greece, Babylonia, and India but transformed into a new systematic form.

Geometry—Derived from the classical Greek postulational formulations with added generalizations and simplification of ideas.

Trigonometry—Developed from the Indian tables of sines with newly added functions and formulas motivated by the Greek geometry of chords.

The Scientific Renaissance

The scholar whose life and works can well represent the period of transition from medieval mathematics to the modern was Johann Muller (1436–1476) of Königsberg, Germany, who was popularly known as Regiomontanus. Regiomontanus graduated from the University of Vienna in 1457 and was appointed as a faculty member, whence he collaborated with his teacher Georg von Peurbach (1423–1469). Regiomontanus accompanied him to Rome to assist in translating Ptolemy's *Almagest*, commissioned by Cardinal Bessarion, which was ultimately bequeathed upon him when Peurbach died within a year. Bessarion was a scholar of mathematics and a native Greek speaker who owned the largest library in Europe. Regiomontanus learned Greek from Bessarion. He lectured in Padua and traveled around northern Italy with Bessarion, where he came in contact with Italian scholars and gained access to the major European libraries. Regiomontanus returned to Germany and settled in Nuremberg, where he wrote his major works under the patronage of a rich merchant. He set up the first scientific printing press, an astronomical observatory, and a small workshop for manufacturing astronomical instruments. Regiomontanus first printed a prospectus to announce his plan to print the translated versions of some important mathematical, astronomical, and geographical works. Then, he started his venture with printing the astronomical textbook *Theoricae Novae Planetarum* (*New Theories of the Planet*) by his teacher Georg Peurbach. Regiomontanus translated the works of Archimedes, Apollonius, Hero of Alexandria, Ptolemy, and Diophantus and made their printed versions available to the general readers. He started working on calendars and built an astrolabe. Regiomontanus also built a sundial for Pope Paul II.

With an objective to provide a systematic trigonometric account to support astronomical studies,

Regiomontanus wrote *de Triangulis Omnimodis* (*On Triangles*). In this treatise of five books, Regiomontanus followed the form of Euclid's writing. The first book included definitions of quantities, ratios, equality, triangles, circles, and other forms, as well as the axioms and theorems he used. The second book covered the solutions of triangles. In the remaining books, he focused on spherical trigonometry in order to strengthen the study of astronomy. These works lack the use of the tangent function, already defined by Nasir Eddin al-Tusi (1201–1274) and the use of fractional decimal expressions, which were to be adopted in a decade's time. His work *Algorithmus Demonstratus* was the first to employ symbolic algebra.

Regiomontanus wrote *Scripta* to describe all the astronomical instruments he collected and/or invented. He used one of them (Jacob's staff) to predict the appearances of a comet for the next 210 years, which was later identified as Halley's comet. He predicted lunar eclipses accurately and used lunar distances to calculate longitudes at sea. This method was also adopted by Christopher Columbus and Amerigo Vespucci in their exploratory voyages.

While working with triangles, Regiomontanus used the algorithmic methods devised by the Arabian algebraists to solve the roots of a quadratic equation of known coefficients in terms of sides of a triangle. This was strikingly different from the ancient Greek works. But owing to the rhetorical form of his expression and his premature death, his successors in Europe had to suffer a delay in the incorporating Arabic methods into the algebra of the ancient Greeks.

The form of mathematical studies of the Renaissance would exhibit features similar to those of the works of Regiomontanus. The explorations of ancient Greek works, the practical examinations of facts and their incorporations in framing the results, scientific collaborations, and publications of the findings to reach out to wider academic circle—all were landmarks in the trajectory of the history of the Renaissance mathematics.

Arithmetic and Algebraic Studies

The best-known exposition of 15th-century Europe was *Summa de Arithmetica, Geometrica, and Proportioni et Proportionalita* of the Italian

mathematician Luca Pacioli (1445–1514). In this book, one can find the first use of the notations p (for +) and m (for –) and the use of the term *cosa* for the unknown, which were echoed throughout Europe in the coming years. The work was a summary of the mathematical knowledge of the time and demonstrated no originality as such, but it still became famous for the illustrious figures drawn in it by Leonardo da Vinci; for the use of algebraic notations instead of rhetorical expressions; and for being the first book to deal with double-entry bookkeeping, games of chance, the golden ratio, rule of 72, and many such things that were to appear much later in the time line of the history of mathematics. Although Pacioli wrongly conjectured, however, that cubic equations cannot be solved algebraically, his work stimulated much thought about the topic.

Scipione de Ferro (1465–1526), a lecturer at the University of Bologna, solved a kind of reduced cubic equation and passed it on to his student Floridus. The University of Bologna at that time was famous throughout Italy for its mathematical competitions. In an effort to win over Floridus in such a competition of solving cubic equations, Niccolo Fontana “Tartaglia” (1500–1557) found solutions to all types of cubic equations, which were eventually published by Heironimo Cardano (1501–1576) in his *Ars magna* and known all over as his (Cardano’s) method of solution of cubics. Lodovico Ferrari (1522–1565), a student of Cardano, influenced by the methods of Tartaglia, worked on the methods of solutions of quartic equations. *Ars Magna* presents the details of solving cubic and quartic equations without any acknowledgment to either of Tartaglia or Ferrari. In attempts of claiming the originality of work, perhaps Tartaglia, Ferrari, and Cardano exhibited the first calculations involving complex numbers (though without much understanding). The controversial character Cardano was a prolific writer and a physician who gained the fame of the ablest algebraist of his time.

Rafael Bombelli (1526–1572), an Italian hydraulic engineer, was keenly interested in algebra. He followed the works of Ferrari, Tartaglia, and Cardan. Bombelli felt that after *Ars Magna*, the learned world needed a careful exposition of algebra, and he wrote *Algebra*. In this book, Bombelli solved many geometric problems algebraically, in a way

similar to that of Regiomontanus, with elaborate geometric illustrations. He designed symbols for the radicals and for the powers of unknown quantities, thereby taking algebra a step closer to modern symbolic representations. Bombelli’s greatest contribution was the identification of conjugates of imaginary numbers and the rules of addition and multiplication of the imaginary numbers. Bombelli later translated a part of Diophantus’s *Arithmetica* and solved many problems, duly acknowledging Diophantus while revising his *Algebra*.

With the availability of the Latin translations of Al-Khwarizmi’s *Algebra*, Arabian algorithm and algebra spread across Europe, and what was called *cosic art* (the study of the unknown) became more and more popular among scholars. In Germany, for example, Johann Widman (1462–1498) published *Behende und hübsche Rechnung auff allen Kauffmanschafften* using the notations of + and – for the first time on print. Rechenmeister Adam Riese (1492–1559) wrote *Die Coss* influenced by the works of Al-Khwarizmi, in which he first replaced the Roman numerals by the Hindu-Arabic numerals. Christoph Rudolff (ca. 1500–ca. 1545) of Vienna, in his book *Coss*, made use of decimal fractions, as well as the modern symbol for roots. But *Arithmetica Integra* of Michael Stifel (ca. 1487–1567) was the most important of all of the 16th-century German algebras. It is outstanding for its treatment of negative numbers, radicals, and powers. In this book, Stifel proposed the use of a single letter to signify the unknown and repeating the letter for higher powers of the unknown. He confidently used negative powers of numbers as well.

Robert Recorde (1510–1558) was the only mathematician of any stature seen in England over a span of almost two and half centuries. He wrote several textbooks to popularize the study of mathematics in his country. He, like Cardano, had a mathematics degree as well as a medical degree and would alternate between the two occupations. His publications *The Grounde of Artes*, *The Pathwaie to Knowledge*, *The Castle of Knowledge*, and *The Whetstone of Witte*, although not outstanding in content, were popular in the absence of other books in the vernacular. The importance of the last book, *The Whetstone of Witte*, lies in the introduction of the = sign. Recorde usually wrote in the medieval question-and-answer style, but the

Pathwaie was different. It was actually an abridged version of Euclid's *Elements*, in which Recorde tried to replace many of the Latin terms by his own verses. But for reasons that remain unexplained, these were not accepted by any successive writers.

Geometry

The 14th and 15th centuries did not witness much advancement in geometry as such. Instead, the geometric models provided the base for developing many important works in art and architecture during the Renaissance. The only remarkable representative in the study of geometry during the Renaissance was Johannes Werner (1468–1522). The work of Werner comprised 22 books entitled *Elements of Conics*. Although it was by no means comparable with the *Conics* of Apollonius (c. 262–c. 190 B.C.E.), it was engaged with the problem of duplicating of cube with a unique and ingenious method quite different from that of Apollonius. Geometrical knowledge was very creatively used by Renaissance painters, sculptors, and architects. It was in this period that a highly realistic linear perspective was developed. Artist Giotto di Bondone (1267–1337) was credited with the introduction of perspective in art. Later Leonardo da Vinci (1442–1519) and his contemporaries included in their works the ideas of ground plane, picture plane, vanishing point, and so forth, and thereby injected the elements of geometric logic to various art forms.

Astronomy and Trigonometry

Astronomy and trigonometry had been important branches of mathematical studies since antiquity. The Renaissance period saw them attain the full bloom in the hands of the Polish astronomer and mathematician Nicholas Copernicus (1473–1543), who gave the revolutionary heliocentric model in *De Revolutionibus Orbium Coelestium*, published in the year he died. Georg Joachim Rheticus (1514–1576) worked thoroughly on Copernican astronomy and wrote explanatory commentaries on his propositions. He also took trigonometry to its present form by writing an elaborate treatise entitled *Opus Palatinum de Triangulis* on the trigonometric functions. Rheticus discarded the forms of the trigonometric functions in terms of

the arcs of a circle and introduced the current definitions of the six trigonometric functions using the sides of a right triangle. He calculated elaborate tables for intervals of 10° in angles. He, however, failed to use the concept of decimal fractions introduced by the Persian scholar Jamshid al Kashi (c. 1380–1429) in the 15th century and used very large numbers for sides of triangles instead. Although essentially the advancement made by Copernicus and his disciple Rheticus marked the beginning of the Scientific Revolution, but as always, revolutions do not wait for a category.

The Scientific Revolution

The period from the publication of *De Revolutionibus Orbium Coelestium* in 1543 by Copernicus to the publication of *Principia* in 1687 by Isaac Newton marks the era called the Scientific Revolution. This revolution changed man's belief and the thought process. The leading personalities of this era tried to explain the relationship of humans with nature from every possible epistemological point of view. From the four qualities that, according to Aristotle, constitute the universe to the existence of the atoms composing everything in the universe, as proposed by Newton, was a great leap in humanity's conceptions of existence. Thinkers of the era overturned the authority of the church, of the physical laws of Aristotle, of the geocentric views of Ptolemy, and of the anatomical and physiological conclusions of the Roman physician Galen that had overshadowed human thought since the beginning of the Common Era.

By the end of the 15th century C.E., there were at least 50 universities across Europe, including the universities founded in the 13th century in Paris, Bologna, Oxford, and Cambridge. Education took an important role in training the youth not only for an occupation but also for a better life. Educationalists included mathematics in all the courses after the reforms of the medieval curricula in the universities. They advocated for the need of mathematics in administration, trade, art, architecture, engineering, and all branches of natural sciences. Some even advocated the need of mathematical studies for spiritual uplift in consonance with the ancient Pythagorean philosophy. Some educationists even integrated the poor into the education

system for the general improvement of society. As a result of all this, mathematics took an important role in the reconstruction of the social world. Many schools were opened, and a new occupation was generated—the *math tutor*—who bore the responsibility of spreading general awareness of mathematics among the masses. Thus, the common people gained access to the knowledge of mathematics as they had never done before.

Astronomy and Geometry

As noted earlier, the introduction of the heliocentric model had revolutionized astronomy and changed the society for good. Galileo Galilei (1564–1642) made fundamental contributions. He claimed that the motion of the planets was actually the motion of falling bodies. His discoveries, pioneering the use of telescope, paved way for the acceptance of the heliocentric model of Copernicus. This simultaneously established experimentation as an important aspect of applied mathematics. Johannes Kepler (1571–1630) took the subject even further by providing laws of planetary motion, which laid the foundation for Newton to establish his universal law of gravitation. In his *Astronomia Nova*, Kepler gave the first two laws of planetary motions.

Kepler also studied the works of Apollonius in detail and came up with the principle of continuity for conics, in which he showed how necessarily moving the focus (possibly to infinity) one can obtain one conic from the other. Kepler also dealt with the problem of the infinitely small. He estimated areas by filling them up with very small triangles, thereby exercising a very preliminary type of integral calculus. The idea of the “points at infinity” as speculated by Kepler was devised formally by Girard Desargues (1591–1661) as the point where the parallel lines intersect. Desargues introduced the perspective theorem and in course developed the subject called *projective geometry*. The ease with 17th-century mathematicians dealt with the concepts of infinity and the infinitesimal completely revolutionized mathematics.

René Descartes (1596–1650) was the first person to associate geometric shapes to algebraic equations in a two-dimensional system (hence named after him as the Cartesian coordinate system) bridging algebra and geometry. Thereby, Descartes constructed the

last segment of the road to infinitesimal calculus. His work in *La Geometrie* makes the first step toward a theory of invariants.

Number Theory

Pierre de Fermat (1601–1665) influenced by *Arithmetica* of Diophantus started working on numbers and their patterns and invented number theory almost single-handedly. He went on to formulate conjectures on the properties of numbers and proving them mostly in communications with friends. He outlined the proofs of some of his conjectures, which his contemporaries often found inadequate in constructing a proof. But later, much of his work was derived from his outlines. The most celebrated conjecture of Fermat was known as the *last theorem*, which remained unsolved for more than 350 years and got tagged to several mathematical academy prizes offered in return for a proof. The theorem finally was proved in 1995, making it the most difficult problem in mathematics to date.

Through communication and collaboration with Fermat, his French contemporary Blaise Pascal (1623–1662) fathered the mathematical theory of probability, which was initiated by Cardano in his expositions on games of chance. Pascal wrote a significant treatise on projective geometry. However, the name of Pascal is attached more popularly to the triangular array of binomial coefficients. Pascal did not construct it rather he derived it from the ancient Greek works. He was the one to provide a proof by defining the recursive relation of the numbers and exhibit the interesting patterns in the triangle, which had come up in many previous works since antiquity.

The Renaissance Function: Logarithm

While working with sequences, the Scottish mathematician John Napier (1550–1617) was once told that the astronomers were overburdened by voluminous calculations in course of their work. This encouraged him to devise some methods of reducing the burden of performing large multiplications and divisions by large numbers. He worked hard and published the *Mirifici Logarithmorum Canonis Descriptio* (*A Description of the Marvellous Rule of Logarithms*) in 1614. The basic idea

behind his invention of logarithms was simplification of large computations. In his work, Napier demonstrated a very interesting kinematic model to explain the formation of the arithmetic and geometric sequences used in the construction of his tables of logarithms. This shows how well equipped the 16th-century mathematicians were to compare arithmetic and geometric sequences and exemplify the result through an example of nonuniform motion. The work of Napier would greatly facilitate the huge calculations of the future astronomers and physicists. Laplace, commenting on the works of Napier two centuries later, said, “By halving the labor of the astronomers he (Napier) doubled their lifetimes.”

Napier built his tables numerically and not geometrically, as the name logarithm (*logos* in Greek means ratios and *arithmos* means numbers) suggests. But he was oblivious of the idea of base, or of the choice of base, which would simplify his calculations further. Upon long conferences with Napier, Henry Briggs (1561–1630), a London mathematics professor, convinced Napier of the use of base 10 instead of his base *e*. Briggs then published his work in 1624 in *Arithmetica Logarithmica* with elaborate logarithm tables of 30,000 natural numbers to 14 decimal places. This function became an indispensable tool of calculations for the next 300 years.

The Renaissance Idea: The Infinitesimal and the Theory of Fluxions

A modest confession of Isaac Newton (1642–1727) was “If I have seen further, it is by standing on the shoulders of giants.” Mastering the geometric ideas of the Greek and the algebraic ideas of Descartes, Newton put forward his theory (published posthumously in *Methodus Fluxionum et Serierum Infinitarum*) of the *fluents* (the variable), the *moments of fluxions* (infinitesimally small increments), and the *fluxions* (flow of the variable). He considered derivatives primarily as velocities and integration as the sum of an infinite series. However, his hesitation to publish the work gave rise to public and private disputes regarding the priority of the discovery of calculus. This was because of Gottfried Wilhelm Leibniz (1646–1716), who published his ideas of tangency, maxima, minima, and estimates of areas in 1684 in the *Acta Eruditorum* (one of

the earliest monthly scientific journals), which brought the theory of calculus to the fore before Newton could do so. Leibniz had studied the works of Descartes and Pascal and was in personal contact with the Dutch mathematician Christiaan Huygens. In his publications, Leibniz gave details of the rules of differentiation and some of the rules of integration. Later, he was joined by the Bernoulli brothers, Jacob and Johann, and they together put forward what we see today as undergraduate calculus. Most of the calculus notations we use today were designed by Leibniz. Both Newton and Leibniz failed, however, to clear the idea of the infinitesimal (the moment of fluxion of Newton and point of tangency of Leibniz), which was attacked heavily by Bishop George Berkeley (1685–1753) calling them as “ghosts of departed quantities.” This inspired further constructive work, and calculus then became invaluable in almost all fields of scientific endeavor.

During the 17th century, several mathematical academies and societies were formed under royal patronage in different parts of Europe. Parallel to the mathematical studies and teachings at the universities, these bodies promoted further research by providing financial assistance, offering prizes and awards, and helping the scholars to publish their research findings in scientific journals all over Europe. These mathematical societies became the conglomerate of the mathematicians collaborating with each other. Enthusiasts like Marin Mersenne and John Collins acted as correspondents, keeping their contemporaries informed about the challenges and open problems of mathematics and inviting their solutions. It is through the exchanges and collaborations of mathematical activities under the supervision of the chosen stalwarts heading such societies that the scholars entered into healthy competitions, taking the subject to soaring heights. The Royal Society of London, French Academy of Science, Italian Royal Society, Royal Academy of Berlin, Royal Academy of Paris, and St. Petersburg Academy all played vital roles in the spreading a strong research culture in mathematics.

Rupa Pal

See also Geometry; Geometry, Non-Euclidean; Mathematics, Antiquity; Mathematics, Middle Ages

Further Readings

- Bell, E. T. (1945). *The development of mathematics*. McGraw Hill.
- Cajori, F. (1909). *A history of mathematics*. Macmillan.
- Eves, H. (1953). *An introduction to the history of mathematics*. Holt, Rinehart & Winston.
- Gregersen, E. (2011). *The Britannica guide to the history of mathematics*. Britannica Educational Publishing.
- Hodgkin, L. (2005). *A history of mathematics: From Mesopotamia to modernity*. Oxford University Press.
- Katz, V. J. (2008). *A history of mathematics: An introduction*. Addison-Wesley Longman.
- Kilne, M. (1972). *Mathematical thought from ancient to modern times*. Oxford University Press.
- Kleiner, I. (2011). *Excursions in the history of mathematics*. Springer.
- Merriman, J. (2010). *A history of modern Europe*. W. W. Norton.
- Merzbach, U. C., & Boyer, C. B. (2011). *A history of mathematics*. John Wiley and Sons.
- Pickover, C. A. (2009). *The math book*. Sterling.
- Plofker, K. (2009). *Mathematics in India*. Hindustan Book Agency.
- Stillwell, J. (2002). *Mathematics and its history*. Springer-Verlag.
- Struik, D. J. (2012). *A concise history of mathematics* (Vol. I and II). Dover Publications.
- Wiener, P. P. (1974). *Dictionary of the history of ideas: Studies of selected pivotal ideas*. Charles Scribner's Sons.

MATHEMATICS, 20TH CENTURY

The 20th century saw the trends of the 19th century continue in terms of the increasing abstraction and generalization of mathematics. The notion of axioms as *self-evident* truths became obsolete as an emphasis on logical concepts such as completeness and consistency took hold. During the 19th century, mathematics also became a growing domain for professional employment, with thousands of new jobs created in teaching and in industry. Hundreds of new fields of mathematics came into play—such specialized areas of study as knot theory, group theory, topology, functional analysis, sheaf theory, singularity theory, graph theory, and others. This entry describes the major developments of the period, focusing on the conceptual breakthroughs and notable achievements of

prominent mathematicians whose work has served to advance the discipline.

G. H. Hardy and Srinivasa Ramanujan

G. H. Hardy (1877–1947) was, along with his protégé Srinivasa Ramanujan (1887–1920), a leading light of early 20th-century mathematics. Hardy's aim was to solve the problems left behind by the previous century, including the Riemann hypothesis. A somewhat eccentric figure, Hardy is probably best known for his work on number theory and mathematical analysis. As a prodigy from a young age, he spent the best part of his academic career at Cambridge University. He is often recognized for reforming British mathematics at the start of the 20th century by instilling a rigor more often associated with that of continental mathematics. Hardy greatly admired German and Swiss mathematicians and was determined to bring to Britain the new tradition of pure, as opposed to applied, mathematics, a type of mathematics which had flourished in Britain since the times of Isaac Newton. He was proud of the fact he had never contributed to anything military or commercial, and he was outspokenly pacifist in his views.

Shortly before the outbreak of the First World War, Hardy claimed to have solved the Riemann hypothesis. What in fact he had accomplished was proof that there were infinitely more zeroes on Riemann's *critical line*. He did not, however, succeed in proving that other zeroes which weren't on *the line* existed—or indeed, that other zeroes elsewhere could not exist.

Hardy famously mentored the talented mathematician Srinivasa Ramanujan, who was originally from Madras in India. Ramanujan wrote to Hardy claiming that he had discovered a formula for calculating the number of primes up to 100 million, a formula that had virtually no errors. Ramanujan was obsessed with proving Riemann's results despite the fact he had had almost no formal mathematical tuition and in fact claimed that most of his ideas came to him in dreams. Hardy, unlike his peers, recognized Ramanujan's genius and invited him to Cambridge University, where he was to befriend and mentor Ramanujan for many years. Hardy and Ramanujan were to collaborate on many mathematical puzzles, although the solution

to the Riemann hypothesis eluded them for their lifetimes. Ramanujan was a prolific mathematician and is believed to have proved or conjectured over 3,000 theorems, equations, and identities, including the partition function, along with its mock theta functions, and the properties of highly composite numbers. He was also involved in work on modular forms, prime number theory, gamma functions, and hypergeometric series problems. Ramanujan was also recognized for his identification of numerous efficient and rapidly converging infinite series, used to calculate the value of π . Some of these series could compute eight decimal places of π with each term in the series. Ramanujan's series are today the basis of the fastest algorithms used by computers, which can now compute to about 5 trillion decimal places.

Ramanujan died at just 32 after falling into a depression and becoming seriously ill. After his death, certain of Ramanujan's other discoveries including the Ramanujan prime and theta function have paved the way for further research, finding practical applications in such varied fields as string theory and crystallography. Hardy, however, lived for a further 27 years, his greatest contribution to mathematics being, in his words, *finding Ramanujan*.

In a related development, the German mathematician Johan Gustav Hermes managed in 1904 to complete construction of a regular polygon with 65,537 sides, using just a compass and straight-edge. This accomplishment took him 10 years.

The early 19th century also heralded the rise of mathematical logic, which built on the earlier work of Gottlob Frege. Mathematicians Giuseppe Peano, David Hilbert, L. E. J. Brouwer, Bertrand Russell, and Alfred North Whitehead took over the mantle, with one of the century's most influential works, *Principia Mathematica*, a joint publication by Russell and Whitehead.

David Hilbert

Right at the beginning of the century, German mathematician David Hilbert (1862–1943) gave an address at the Sorbonne, Paris, in which he set out what were in his view the 23 most important unsolved mathematical problems of the times. *Hilbert's problems*, as they became known, set the agenda for many decades to come. Of the 23

problems, 10 have thus far been solved, seven partially solved, and two, the ever-perplexing Riemann hypothesis and the Kronecker–Weber abelian theorem, still unresolved. Four of the problems are considered to be too vaguely formulated for them to be embarked on.

Hilbert, like many other mathematicians before him, had attended Göttingen University and spent much of his working life there. Considered an *out of the box* thinker, Hilbert went on to have numerous mathematical terms named after him, for example, *Hilbert space*, *Hilbert curves*, and *Hilbert inequality*. He also devised some important theorems, eventually becoming one of the best-known mathematicians of the 20th century. In his younger days, Hilbert had worked on number theory and abstract algebra, and then, he changed direction to work on integral equations. His work on integral equations helped to reinvent the field and led to what was known as his *finiteness theorem*. Hilbert managed to prove that, while there were an infinite number of possible equations, it was nonetheless possible to split them into a finite number of *kinds* of equations. These could then be used to build all other equations. Although Hilbert himself was unable to build these equations, he could simply prove the necessity of their existence. This type of proof is referred to as an *existence proof* rather than a *constructive proof*. Critics at that time tended to dismiss Hilbert's proof as having no practical basis, but his proof was to set the tone for a brand-new style of abstract mathematics. The use of existence proofs rather than constructive proofs became an implicit aspect of his development of Hilbert space. This was a generalized form of Euclidean space, an extension of the use of vector algebra and calculus to spaces possessing a finite or possibly even infinite number of dimensions. This progression of space theory formed the basis of much of the contributions to mathematical physics for many years to come. Some consider it to be one of the best explanations of quantum mechanics in existence.

When giving his famous enunciation of the problems of mathematics, Hilbert was convinced that solutions would be found, even postulating that there are no mathematical problems that cannot be solved. "We must know! We will know!" was famously inscribed on his tombstone as

evidence of his conviction. His belief that mathematics should be put on logical foundations where there were no limits was behind his unshakeable optimism. For Hilbert, the basis of mathematics was not in logic per se, but in a simpler system of prelogical symbols. These could be amassed to form strings and axioms, then a set of rules, or *rules of inference*, could be applied to them. He devoted much time to finding a consistent, complete set of axioms for the whole field of mathematics (this was *Hilbert's program*). By the start of the 20th century, Hilbert had set out an entirely new formal set of geometrical axioms, which became known as *Hilbert's axioms*, as opposed to the traditional axioms devised by Euclid.

This work, however, was definitively stopped in its tracks by Kurt Gödel's theory of incompleteness in the 1930s. The desire to understand Gödel's work however furthered a new objective, that of developing recursion theory and mathematical logic as an autonomous discipline, something that was later to provide a foundation for theoretical computer science.

Hilbert could only look on in horror at the ensuing Nazi repression of his Jewish mathematician friends and colleagues in Germany during the 1930s as well as the systematic destruction of what was the one of the foremost mathematical centers in human history. At the time of his death in 1943, Hilbert's work was little known, and he died in obscurity.

Bertrand Russell and Alfred North Whitehead

Two other towering figures in the 20th century mathematics were Bertrand Russell and Alfred North Whitehead, mentioned earlier in this entry. Both were mathematicians, philosophers, and logicians, who disparaged *continental idealism*. Both made significant contributions to the fields of mathematics, logic, and set theory.

Russell is considered one of the chief founders of analytic philosophy, while his mathematics was heavily influenced by the set theory and logic of Gottlob Frege—Frege having developed the earlier work of Georg Cantor. Russell is well known for his *paradox* of a set containing all of the sets that are not members of themselves. This indicated that Frege's *naive* set theory entailed contradictions.

The paradox is often illustrated using this example: "If a barber shaves only and all the men in the village who do not shave themselves, does he shave himself?" Russell's paradox appeared to put the whole of mathematics on a shaky footing with the implication that perhaps not even mathematics could be relied upon to provide us with the truth in absolute terms. Russell's work certainly shook Frege's confidence, something he was honest enough to admit in an appendix to his *Basic Laws of Arithmetic*.

But it was the *Principia Mathematica* that was to seal Russell's fame, a work published in three volumes between 1910 and 1913. The first he cowrote with Whitehead, but the latter two were written almost entirely by Russell alone. The *Principia* was an attempt to derive all mathematics from a set of purely logical axioms without running into the kinds of paradoxes that plagued Frege's work on set theory. Russell succeeded in this by devising a system of *types*, whereby objects of a certain type are built exclusively from objects of types which precede them, lower down the hierarchy. This ensures that the problem of loops is avoided. Since each set of elements is of a different type from each of its elements, we cannot view the system in terms of *sets of sets* and we thereby avoid the usual paradoxes. The *Principia* also required the addition of three further axioms, which were true not simply because of the laws of logic. These were, first, the *axiom of infinity* to guarantee the existence of at least one infinite set—the set of all natural numbers; second, the *axiom of choice* to ensure that given any collection of *bins* containing at least one object, it is possible to select exactly one object from each bin. This must be possible even if there are an infinite number of bins and although there is no set *rule* about which object to choose from each bin. The third axiom was Russell's *axiom of reducibility*, which would ensure that propositional truth functions can be expressed by predicative truth functions.

The extensive scope of the *Principia* is illustrated by the fact that it takes 360 pages to prove that $1 + 1 = 2$. The book is still considered one of the most important works in the history of mathematics, on a par with Aristotle's *Organon*. In the 1930s, Gödel's incompleteness theorem struck again, however, proving that the *Principia's* contents were neither consistent nor complete.

Kurt Gödel

Kurt Gödel (1906–1978) was an Austrian mathematician, who put intense constraints on what could and could not be proven. He effectively turned mathematics on its head with his incompleteness theorems. They seemed to prove what had hitherto been unthinkable—that there were solutions to mathematical problems that could not be proven—ever. Gödel studied number theory at the University of Vienna but became fascinated by mathematical logic, and it was to this he then devoted his entire career. Early in his life, he, like Hilbert before him, was optimistic that mathematics could be made whole and unproblematic. He became part of the famous Vienna circle of intellectuals during the First World War and discussed philosophical ideas with logical positivists such as Rudolf Carnap, who rejected metaphysics as empty and meaningless, in favor of a codified basis for the language of science. Gödel was influenced by this positivistic outlook although he himself did not become a logical positivist. Instead, he turned his attention to finding a logical foundation for all mathematics. What he discovered, however, was that there was no such thing.

In recognition of Hilbert's optimism that there could be a logical basis for mathematics, Gödel first drew up his completeness theorem. This illustrated that all the valid statements in Frege's *first-order logic* can be proven from a set of basic axioms. When it came to extending that logic to include arithmetic, however, the difficulties appeared. Gödel's incompleteness theorem (or theorems, since there were actually two) showed that in any mathematical system at least that powerful, there would be statements that, although they might be true, could never be proven. Gödel had used the plain language assertion "this statement cannot be proved," which was an updated version of the *liar paradox*, an assertion which of itself must be true or false. If the statement is false, it shows that the statement can be proved, thereby suggesting that it is in fact true—and thus generating a contradiction.

In order for Gödel to apply this deduction to mathematics, he had to convert the statement into a formal language or pure statement of arithmetic. He did this by the use of a code based

on prime numbers, in which strings of prime numbers take over the role of natural numbers, rules of grammar, operators, and all other aspects of formal language. The mathematical statement that ensued appeared just like its natural language equivalent as true, but impossible to prove.

The incompleteness theorem caused a crisis in the world of mathematics. To be able to state something as true but be unable to prove, it was a mathematician's worst nightmare. Gödel thereby stopped many mathematicians in their tracks, including Russell and Hilbert. He had provided proof that any system of numbers or logic dreamt up by a mathematician would fail to succeed in coming up with a full set of provable assumptions.

Gödel, like Cantor before him, also applied himself to the problem of the *continuum hypothesis*. He made an important development in resolving it by proving that the axiom of choice is independent from finite type theory. Unfortunately, Gödel, like Cantor, also suffered from a form of mental deterioration and underwent a series of mental breakdowns. He was also forced to witness the destruction of his mathematical community at the hands of the Nazis, which caused him to flee to the United States. At Princeton University, he befriended Albert Einstein and made contributions to Einstein's theory of general relativity. Gödel is recognized as one of the century's most influential mathematicians; however, there were some who found his views untenable. The majority of mathematicians at the very least found it hard to come to terms with Gödel's conclusions.

Alan Turing

One mathematician who found it extremely hard to countenance Gödel's conclusions was British mathematician Alan Turing (1912–1954). Famous for his work during the Second World War on code-breaking at Bletchley Park, Turing's work led to the breaking of the Germans' Enigma code. Turing was also responsible for taking Gödel's bleak theorem even further. It is this, and for his large volume of work on computer science, that Turing is known for.

Turing directed much effort into the clarification and simplification of Gödel's theory of

incompleteness. His intention was to make it far less abstract and more concrete, and his solution, which he later claimed had come to him in a dream, would lead to one of the most significant inventions of the modern world, the computer.

Turing repositioned incompleteness in terms of computers, or at that time what were seen as theoretical devices that could manipulate symbols; these devices came to be called *Turing machines*. He replaced Gödel's universal arithmetic-based language with his formal, simple device and illustrated that this device could perform any conceivable computation that was representable as an algorithm. Next, he managed to prove that, although this highly logical machine was driven by arithmetic, there would always be certain problems it could never solve. When an unsolvable problem was fed into the computer, it would carry on attempting to solve it but would never succeed. Determining the general conditions for when this would happen was called the *halting problem*. Turing also discovered that we could not know in advance which problems were unprovable, thereby providing a negative proof of David Hilbert's *decision problem*. After the war, Turing worked on developing computers, such as the ACE or automatic computing engine, and the Manchester Mark 1. Despite the fact that computer technology was in its infancy, Turing was convinced of its potential. He foresaw a future where computers could think, play games, and communicate, and he was the first to develop a design for computer chess. Turing also was the first to look at the notion of artificial intelligence. The *Turing test* was developed by him to define the point at which a computer could be considered *intelligent*. This involved the computer's being able to fool a human into thinking it was also human. Unlike Gödel, who believed in the power of intuition and was convinced that the human mind had capabilities outside of the systems he described, Turing preferred to think of the mind as a formalized system such as the one embodied by a computer.

John von Neumann

John von Neumann (1903–1957), another child prodigy, became one of the foremost mathematicians of the 20th century. Neumann made huge

contributions to many fields of mathematics and also worked in quantum physics and on the Manhattan Project, which led to the development of nuclear physics. He is well known for his involvement in game theory and for his design for a digital computer that could hold both a processing unit and a storage structure that could hold instructions and data. This general plan is one that most computers still use today.

Andre Weil

The French mathematician Andre Weil (1906–1998) was active chiefly in the mid part of the 20th century. He made significant contributions to many areas of mathematics and was in particular interested in identifying the connections between algebraic geometry and number theory. He became fascinated by Diophantine equations, and it was this that led to his first notable work in mathematics, on the theory of algebraic curves. He went on to discover the topological algebraic adèle ring, which he built using rational numbers. He was also a founder and early leader of the Bourbaki group of mathematicians in France, a group that published many influential textbooks on advanced mathematical topics and whose aim was to unify mathematics via set theory. At the outbreak of the Second World War, Weil, a conscientious objector, fled to Finland. Later making his way back to France, he was arrested and imprisoned for refusing to perform military service. In 1940 while in prison in Rouen, Weil carried out work that sealed his reputation in mathematics. Building on the work of Évariste Galois in the 19th century, Weil decided to use geometry to analyze equations, thus developing algebraic geometry, which offered up a new language for understanding the solutions to equations. Weil's work on zeta-functions succeeded in proving the Riemann hypothesis for curves over finite fields, thereby laying the foundations for abstract algebraic geometry, the modern theory of abelian varieties, and the theory of modular forms. Weil's work on curves has influenced many areas of mathematics as well as other fields including string theory and elementary particle physics.

One of Weil's last contributions to the world of mathematics was the Tamagawa numbers, although his conjectures were not to be proven until 1989.

He also played a big part in the formulation of the *Shimura–Taniyama–Weil theory* on elliptic curves, later taken up by Andrew Wiles as part of his proof of Fermat’s last theorem. His development of what is known as the *Weil representation* later provided a contemporary framework for understanding classical quadratic forms.

Alexander Grothendieck

One of Weil’s greatest legacies was in his heir apparent, Alexander Grothendieck (1928–2014), a colorful figure in 20th-century mathematics. Grothendieck was a structuralist, interested in the structures that lie beneath the surface of mathematics. During the 1950s, he devised a new and powerful language that allowed mathematical structures to be seen in a new way. This meant that new solutions in number theory and geometry and even in basic physics could be discovered. Grothendieck’s *theory of schemes* was instrumental in allowing for Weil’s number theory conjectures to be solved, and his *theory of topoi* assumed an important place in the field of mathematical logic.

Grothendieck’s early work was in functional analysis and key contributions included topological tensor products of topological vector spaces, the application of L spaces in the study of linear maps, and the theory of nuclear spaces. Working between the years 1949 and 1953, Grothendieck soon became a leading figure in the world of functional analysis. It was, however, in algebraic geometry, where he was to complete his most influential work. Around 1955, he began work on *sheaf theory* and *homological algebra*, producing his influential paper “Sur Quelques Points de l’Algebre Homologique” in 1957. In this article, he introduced the ideas of abelian categories and used them to illustrate that sheaf cohomology is definable in terms of certain derived functors. Homological methods and sheaf theory had been introduced by Jean-Pierre Serre, among others, earlier in the century and sheaves themselves were defined by Jean Leray. Grothendieck was able to take them to a higher abstract level than before, turning them into key principles of his theory. He switched the focus from studying individual varieties to pairs of varieties related by a morphism, a shift that enabled a broad generalization of numerous classical theorems. The first such

theorem was the relative version of Serre’s theorem, where it was shown that the cohomology of a coherent sheaf is not infinitely dimensional.

Grothendieck also used the same thinking when he came up with an algebraic proof for the Riemann–Roch theorem, which was his first major achievement in algebraic geometry, as well as an algebraic definition of the fundamental group of a curve.

Grothendieck’s work in algebraic geometry was revolutionary in much the same way as Cantor’s, Gödel’s, and Hilbert’s, which is why he is considered by many to be one of the driving mathematical forces of the 20th century. Grothendieck was admired for his mastery of abstract principles and his perfectionism in formulating and presentation. His influence extended into many branches of mathematics, for example, the modern theory of D modules. He produced relatively little output after 1960 and retreated from public life to live in relative estrangement from the mathematical community.

Paul Erdős

Paul Erdős (1913–1996) stood out as a prolific albeit eccentric figure in the world of mathematics during the 20th century. Erdős hailed from Hungary and collaborated with many other mathematicians on various problems including graph theory, number theory, combinations, approximation theory, probability theory, and set theory. He became well known for offering prizes for solving unresolved mathematical problems including his own conjecture on arithmetic progressions, many of which remained unsolved in his lifetime. Erdős came from a Jewish family and was born in Budapest just before the outbreak of World War I.

Erdős was a legendary mathematician who remained devoted to his subject until he was into his 80s. He was one of the century’s most prolific mathematicians and produced more than 1,500 works during his lifetime. At age 20, he first made his mark in the field when he discovered a proof for a famous theorem in number theory. The theorem, Chebyshev’s theorem, states that for each number greater than one, there is always at least one prime number between it and its double. A prime number is a number that has no divisors other than itself and 1.

Erdős kept his interest in number theory for the whole of his life, posing and solving problems that

involved relationships between numbers; although they may have looked simple, these problems were very difficult to solve. Erdős believed that mathematical truths are discovered rather than invented, and he had a characteristic way of displaying this notion. He spoke of a Great Book in the Sky, maintained by God, where all the most elegant mathematical proofs were kept.

Paul Joseph Cohen

On the other side of the Atlantic, Paul Joseph Cohen (1934–2007) was a mathematician benefiting from the influx of European talent to the United States during the war years. He had a broad range of mathematical interests, from mathematical analysis to differential equations, number theory, to logic. In the 1960s, Cohen applied himself to Hilbert's list of problems yet to be solved, including Cantor's continuum hypothesis. This questioned whether there exists a number of numbers bigger than the set of all natural numbers yet smaller than the set of real numbers. Cantor had been convinced that there was not, but failed to prove his view, a view that had not been satisfactorily proven since then, although some progress had been made. Ernst Zermelo and Abraham Fraenkel had come up with the standard form of axiomatic set theory in the early 1920s. Termed the *Zermelo–Fraenkel* (ZF) *set theory*, this went on to become the basis for much of mathematics thereafter. Kurt Gödel showed in 1940 that the continuum hypothesis, because it is consistent with the ZF theory, cannot be disproved by it. This is the case even if the axiom of choice is selected (otherwise known as the *ZFC theory*). Cohen wanted to show that the continuum hypothesis was or was not independent of the ZFC theory, which would prove that the hypothesis was independent from the axiom of choice.

Cohen arrived at a startling conclusion, based on a new technique termed *forcing*, which he had developed himself. Cohen deduced that both answers could be true; in other words, that the continuum hypothesis and the axiom of choice were separate from ZFC set theory. From this, it became theoretically possible for there to be two different mathematics, which were internally consistent. One set of mathematics would see the continuum hypothesis as true and where there

would not be any set of numbers, and another where the hypothesis was false and a set of numbers existed. The proof looked to be correct, but because Cohen's methods were so novel, there was no consensus of opinion until it was finally given the stamp of approval by Gödel in 1963.

Cohen's findings were revolutionary, and ever since mathematicians have constructed two mathematical worlds in which the continuum hypothesis either applies or does not. Modern mathematics demands in fact that proofs must state whether the result is dependent or not on the continuum hypothesis. Cohen's proof was so successful that he decided to tackle something even more difficult, the Riemann hypothesis, and devoted more than 40 years to the problem, right up until his death in 2007. The problem still has no solution, however, although Cohen did give hope to other mathematicians, including some of his brightest students.

Julia Robinson and Yuri Matiyasevich

In the 20th century, several women, including Alicia Stout and Emmeline Noether made significant impacts on mathematics, but perhaps the most prestigious was Julia Robinson (1919–1985). Robinson spent much of her mathematical career in attempting to determine computability and decision problems, the questions that, in formal systems, have a *yes* or *no* answer, the answers being dependent on the parameters that had been input. She was also dedicated to answering one of Hilbert's mathematical problems, that of ascertaining whether it was possible to discover whether or not any particular Diophantine equation had a whole-number solution. There was a growing feeling that this was impossible—but neither was there a way to prove that it could never be possible.

In the 1950s and 1960s, Robinson, together with collaborators Hilary Putnam and Martin Davis, pursued the problem, eventually developing what was known as the *Robinson hypothesis*. This proposed that by constructing one equation with a solution of very specific set of numbers, of which one grew exponentially, it could be proved that no such method existed. Robinson later enlisted the help of Russian mathematician Yuri Matiyasevich (b. 1947) to help her solve the problem in its entirety. Matiyasevich was something of

a mathematical prodigy, having already won several prizes for his work. He used Hilbert's 10th problem as the basis for his doctoral thesis and had started to correspond with Robinson about her work on the problem. Matiyasevich eventually identified the missing piece of the puzzle in 1970 at the age of 22. He found a way to capture the Fibonacci number sequence using equations that formed the basis of the Hilbert problem. By building on Robinson's work, it could finally be proved that it is impossible by any method to determine in a finite number of processes whether Diophantine equations can be solved in rational integers. Despite the cold war between the United States and the Soviet Union being at its height, Matiyasevich was quick to acknowledge Robinson's input into his findings, and they continued to work together up until Robinson's death in 1985. Matiyasevich also made developments into prime number theory, devising a kind of *visual sieve* that crosses out composite numbers and leaves behind only prime numbers. His theorem on recursively enumerable sets is named after him, and he is credited with a polynomial related to the triangulation of spheres.

The 1970s and Beyond

The field of complex dynamics was taken up by the French mathematicians Pierre Fatou and Gaston Julia in the early part of the 20th century. However, it was not until the 1970s and 1980s that the field gained momentum. In addition to complex dynamics, chaos theory benefited from the ongoing development of computers, with their ability to carry out repeated iterations of mathematical formulas on a huge scale. Chaos theory deduces that certain systems appear to act randomly even when they are not actually random. Conversely, some systems can behave perfectly predictably in many ways but are unpredictable when it comes to detail. A chaotic system can be mapped graphically to understand its possible behaviors, and these mappings, or *strange attractors*, are by nature fractal. When the observer zooms in to get to more detail, the overall pattern stays the same at that magnification. Edward Lorenz was an early proponent of chaos theory. His interest in the subject came about accidentally because of his work in predicting the weather.

A computer model he was running was accidentally saved using three-digit numbers instead of the six digits he had been using. This small rounding error led to a wide variation in the results. Lorenz found that tiny changes in the initial conditions can produce big changes in long-term outcomes. He termed this effect the *butterfly effect* and managed to demonstrate this event using his Lorenz attractor, a fractal structure that bore correspondence to way the Lorenz oscillator behaved. The Lorenz oscillator is a three-dimensional system that illustrates chaotic flow.

In 1976, Kenneth Appel and Wolfgang Haken devised the first proof of the *four-color theorem*, which was the first major theorem proven by a computer. The four-color theorem had initially been the work of the 19th-century mathematician Francis Guthrie. It proposed that in a given segregation of a plane into continuous regions (called a *map*), the different regions require no more than four colors in order to color them, so that no two adjacent regions has the same color. Alfred Kempe provided a proof for this in 1879; however, this was disproved 11 years later by Percy Heawood, when he provided his five-color theorem. Appel and Haken's solution took more than 1,200 hours of computer time and the examination of approximately 1,500 configurations.

The 1970s also saw the increasingly popular use of origami as a mathematical method. This form of paper folding, which is usually associated with hobbyists, became as powerful as Euclidean geometry in certain ways. Margherita Piazzolo Beloch had shown, in 1936, how paper could be folded to get the cube root of its length. It was not until 1980, however, that it was proved that the origami method could be used to solve the problem known as *doubling the cube*. This puzzle had defeated mathematicians as far back as the ancient Greeks. This was followed in 1986 by a demonstration of how origami could solve an equally difficult problem—that of *trisecting the angle*. Japanese origami expert Kazuo Haga's folding techniques have derived many surprising results and won him the honor of having several mathematical theorems named after him.

In 1995, the British mathematician Andrew Wiles solved Fermat's Last Theorem, just 350 years after Fermat had posed the question. Wiles had been determined to solve this theorem from

an early age, and he refused to be defeated. He was not, however, solely responsible for proving the theorem and collaborated with many other mathematicians including Gerhard Frey, J. Pierre Serre, Yutaka Taniyama, Ken Ribet, and Goro Shimura. Wiles did provide the links and final synthesis and the eventual proof for semistable curves—a proof that runs more than 100 pages in length.

In recent years, another important conjecture, the Poincaré Conjecture, was also solved. In 2002, Grigori Perelman managed to provide proof of the thesis, which was then more than 100 years old. Perelman turned down the \$1 million prize, however, stating that for him, if his proof was correct, he needed no other recognition. The conjecture turned proof states that if a loop situated in a connected, finite, three-dimensional boundary-less space that can be tightened continuously to a point just like a loop drawn on a two-dimensional sphere, then the space is a three-dimensional sphere. Perelman's elegant but complex solution involved looking at the three ways, in which three-dimensional spheres can be enclosed in higher dimensions. He also contributed significantly to topology and Riemann's geometry.

John Nash was an American economist and mathematician who, although he suffered from paranoid schizophrenia, completed important work in differential geometry and in game theory. He also worked on partial differential equations, something which helped to illuminate the factors behind chance and events in such areas of daily life as computing, accounting, and military theory.

Nash earned a doctorate for his work on non-cooperative games, a dissertation that contained the definition and properties of what would be called the *Nash equilibrium*.

British mathematician John Horton Conway is responsible for bringing in the "Game of Life" rules in 1970. These principles that govern the pattern of cells evolving and growing in a grid are an example of *cellular automation*, and they were taken up by computer scientists. Conway also added his input to pure mathematics including group theory, geometry, number theory, and game theory. He is known for his invention of strange-sounding concepts including *monstrous moonshine*, *the grand antiprism*, and *surreal numbers*. He is also the creator of such games as Philosopher's Football and the Soma Cube. Rubik's Cube is another well-known

mathematical-based puzzle developed in the 20th century, which became incredibly popular with the general public. Rubik's Cube came into existence in 1974 and was the brainchild of Ernő Rubik, a Hungarian sculptor and professor of architecture. The puzzle, he predicted, would never be solved in his lifetime if it were only randomly twisted, so he began working on a solution, beginning with the alignment of the eight corner cubes. He then had to work out a sequence of moves for rearranging just a few cubes at a time and within a month he had solved the puzzle.

Sudoku was another mathematical puzzle that proved popular with the public and also generated serious attention from mathematicians, who sought to explore the theoretical limits of the game. The game of Sudoku is believed to have had its roots in the mathematical concept of Latin Squares. The Swiss mathematician Leonhard Euler developed a system in the 1780s that involved arranging numbers in such a way that any number or symbol would occur only once in each row or column. Latin Squares are actually used in statistical analysis.

The development and global dissemination of electronic computers in the 20th century have helped immensely in identifying phenomena such as Mersenne prime numbers. Mersenne primes are the numbers of less than a power of two. In 1952, a computer that was known as SWAC identified the 13th Mersenne prime number. This was the first to be discovered in more than 75 years and in time computers went on to identify several large numbers.

The internet has also brought with it a host of mathematical possibilities. In the 1990s, the Great Internet Mersenne Prime Search increased the rate at which Mersenne primes could be found. Run by volunteers and using freely available software, this has advanced research considerably. The 45th Mersenne prime was discovered in 2009 and was a landmark discovery, since it was also the largest prime number ever found, containing nearly 13 million digits. The search continues too for more and more accurate approximations of the irrational number π , a number that stands at 5 trillion decimal places and still counting. Another problem that computers have helped to advance is the P versus NP problem. This was introduced in 1971 by Stephen Cook and became a notoriously

complex unsolved theory. In its most basic form, the problem asks whether every problem whose solution can be checked efficiently by a computer can also be efficiently solved by a computer. That is, whether there are questions whose answer can be rapidly checked but which require an unending amount of time to solve by any kind of direct process or procedure. Although this problem, known as *Cook's Theorem*, or the *Cook–Levin Theorem*, may sound simple enough, computer scientists and mathematicians have struggled to find the answer to it for over 40 years. One possible solution, put forward by Vinay Deolalikar in 2010, purports to prove that P is not equal to NP (therefore, problems that are nonsolvable yet easily checked exist) has not as yet been accepted as being correct.

Finally, mathematics as a subject is clearly set to grow ever larger with computers becoming more important and powerful than ever before and the application of mathematics to bioinformatics growing exponentially. The end of the 20th century saw an explosion in knowledge, and there are currently hundreds of specialized areas in mathematics, and ever more mathematical journals published both in print and online. The 21st century has already seen the Clay Mathematics Institute announce the seven Millennium Prize Problems, and the Poincaré conjecture finally solved by Grigori Perelman.

Thora L. Fitzpatrick

See also Mathematics, 19th Century; Mathematics, Antiquity; Mathematics, Enlightenment; Mathematics, Middle Ages.

Further Readings

- Bradley, M. J. (2006). *Mathematics frontiers: 1950 to the present* (Pioneers in Mathematics). Facts on File.
- Clapham, C., & Nicholson, J. (Eds.). (2014). *The concise Oxford dictionary of mathematics*. Oxford University Press.
- Gray, J., & Rowe, D. (2000). *The Hilbert challenge: A perspective on twentieth century mathematics*. Oxford University Press.
- Hoffman, P. (1998). *The man who loved only numbers: The story of Paul Erdos and the search for mathematical truth*. Hyperion.
- Siegmund-Shultze, R. (2009). *Mathematicians fleeing from Nazi Germany: Individual fates and global impact*. Princeton University Press.

- Schaaf, W. (1969). *Memorable personalities in mathematics: Twentieth century*. Stanford University Press.
- Segal, S. J. (2003). *Mathematicians under the Nazis*. Princeton University Press.
- Mason, F. (2014). *Turing: The tragic life of Alan Turing*. Create Space Independent Publishing Platform.
- Yandell, B. (2003). *The honors class: Hilbert's problems and their solvers*. A. K Peters/CRC Press.

MEASUREMENT

Measurement is widely performed to produce information about properties of objects, and the quality of measurement results may be evaluated: A theory of measurement provides criteria to distinguish measurements from processes like opinion making, and about the quality of measurement results. Moreover, it establishes the structure—whether quantitative or not only, and the domain, whether physical or not only—of properties to be measured.

The Scope of a Theory of Measurement

Measurements are performed in many different contexts to produce information about properties of systems, phenomena, processes, individuals, organizations, and so on. It is an integral part of modern science as well as of engineering, commerce, and daily life. Measurement is often considered a hallmark of the scientific enterprise and a privileged source of knowledge. But what is the source of the acknowledged *special efficacy* of measurement?

First, a theory of measurement is expected to provide a criterion of *demarcation*, able to distinguish measurement from other information-related processes like opinion making and computation: not all processes of information production are measurements. Second, the fact that measurement is a designed-on-purpose process, not a natural phenomenon, adds a pragmatic layer to the theory: A theory of measurement is expected to provide some criteria of *qualification*, able to characterize good and bad measurements (the quality of an opinion, for example, by a highly qualified expert, could be better than the

quality of a measurement as performed, for example, by means of a cheap instrument), and this prevents the defining of measurement simply as any high-quality information production process.

Moreover, a theory of measurement needs to be situated with respect to two strategic issues. First, is there any constraint about the *structure* of measurable properties? And therefore, in particular: is—for some reason to be identified—measurability a feature of quantitative properties (*quantities* for short) only, or can nonquantitative properties also be measured? Second, is there any constraint about the *domain* of measurable properties? And therefore, in particular: Is—for some reason to be identified—measurability a feature of physical properties only, or can psychosocial properties also be measured?

These issues probe the very idea of what measurement is, as developed across millennia of scientific, technological, philosophical, and societal changes, and as currently challenged by the widespread emphasis on datafication: Can an entirely informational process, with no empirical stages in it, be a measurement (perhaps a *virtual measurement*, as sometimes it is called)? Does a difference remain between a measurement and a simulation of a measurement?

In what follows, after some notes setting forth a fundamental terminological ground, the domain, whether physical or not only, and the structure, whether quantitative or not only, of measurable properties are analyzed in historical perspective, and on this basis, the problems of demarcation and qualification are discussed. The conclusion will be that what is measurement, and therefore what a theory of measurement should deal with, is still a moving target.

Background

Measurement is about attributing values to properties of objects: unpacking this is a precondition for a theory of measurement. The assumption is that *objects*, like systems, phenomena, processes, individuals, and organizations, *have properties*, like length, color, age, and intelligence (in a property-oriented ontology an object is something that has properties; whether objects can be identified independently of their having properties is an issue preliminary to our subject). Not any object

o has any property *P* (e.g., gases and companies do not have geometric shape), so that for each property the set of objects is partitioned into the subset of the objects that have the property—it may be called the *domain* of the property, $D(P)$ —or into the complementary subset of the objects that do not have it. This may be written in the predicative form $has_P(o)$, that is either true or false, so that

$$D(P) = \{o : has_P(o)\}$$

Hence, $has_electric_charge(electron)$ is true and $has_reading_comprehension_ability(Moon)$ is false. Such a classification—for any sufficiently well-identified *P* and *o*, either $has_P(o)$ or not—is the piece of information that provides the fundamental justification of the interest in dealing with properties.

Sometimes this information may be refined, by recognizing that a property *P* is associated with a set X_P of other properties, $X_{P,1}, X_{P,2}, \dots$, such that, for each *o*, $has_P(o)$ if and only if there exists one and only one $X_{P,i}$ in X_P such that $has_X_{P,i}(o)$. For example, for a physical body *o*, $has_color(o)$ if and only if $has_color_red(o)$, or $has_color_blue(o)$, and so on. In other words, for having color an object must have *a* color. Such coordinated—because exhaustive and mutually exclusive—properties $X_{P,i}$ may be called *manifestations* of *P*, or *determinates* of the *determinable* *P*. In this way, the domain of *P* is partitioned into equivalence classes, where each $X_{P,i}$ identifies the class of objects *o* such that $has_X_{P,i}(o)$. This process may be repeated by refining the partition: has_color_red can be considered a determinable in turn, whose determinates are $has_color_red_magenta$, $has_color_red_scarlet$, and so on. Note that multiple criteria of partitioning are usually possible, and therefore that the set of determinates of a given determinable is not necessarily unique.

Whenever $X_{P,i}$ is considered as an element of X_P , the predicate $has_X_{P,i}(o)$ is customarily written in the functional form of a *basic evaluation equation*

$$P(o) = X_{P,i} \text{ in } X_P$$

(e.g., $color(o) = red$ in $\{red, blue, \dots\}$), thus highlighting the conditions of existence and uniqueness in X_P of the determinate $X_{P,i}$ of the determinable *P* for each object *o* in $D(P)$. This

equation formalizes the core information that can be obtained by a measurement: the object o has the property P , and in particular the property $X_{p,i}$ in the classification of P according to X_p (e.g., o has *color*, and in particular has *color red* in the classification of *color* according to $\{red, blue, \dots\}$). This justifies the usual understanding of P as a function that when applied to the argument o has the value $X_{p,i}$ in the set X_p and provides a sufficient explanation of the somewhat confused related lexicon: Both P and $P(o)$ are considered to be properties (color is a property; the color of this object is a property that the object has), and $X_{p,i}$ in X_p is called a *value of a property*.

Measurement as Representation

The idea of *measurement as representation* plausibly originates from here: As it was stated by Norman Campbell, who devoted to measurement a large part of his book *Physics: the Elements* (1920), measurement is the process of assigning numbers to *represent* properties, where in fact *representation* and *assignment* are usual terms to designate functions. In this way, the content of the basic evaluation equation is weakened, by omitting any condition about the *equality* of properties of objects and values of properties, and actually interpreting such equations as

$$\text{representation}(P(o)) = \text{value}$$

(or sometimes even as $\text{representation}(o) = \text{value}$, by thus absorbing the reference to the property in the representation). In this broad sense, measurement is considered to be the representation of properties of objects by means of values that are assigned according to some consistency conditions, the simplest of them being that distinguishable properties must be represented by distinguishable values, so that no information is lost in the representation. A basic set of such conditions was proposed by Stanley Stevens in his seminal paper *On the Theory of Scales of Measurement* (1946), in which a sequence of scale types is identified, each of them characterized by an algebraic structure and a related condition of invariance. For example, if the properties x_i to be represented are not only distinguishable from each other but also ordered, then the values $f(x_i)$ by

which the properties are represented must be ordered accordingly:

$$\text{if } x_1 < x_2 \text{ then } f(x_1) < f(x_2)$$

As a consequence, any further monotonic function g applied to the values preserves the information conveyed in the representation; that is,

$$\begin{aligned} &\text{if } x_1 < x_2 \text{ and } f(x_1) < f(x_2) \\ &\text{then also } g(f(x_1)) < g(f(x_2)) \end{aligned}$$

so that the property x can be represented not only by $f(x)$, but also by $g(f(x))$. For this reason g is called an admissible scale transformation.

Many physical properties are more than ordinal only and admit a comparison by ratio (e.g., this is twice as much as that), sometimes based on an additive structure, such that objects can be composed according to their property and the composition has the features of a sum: $x_1 + x_2 = x_2 + x_1$, $(x_1 + x_2) + x_3 = x_1 + (x_2 + x_3)$, and so on. At least in classical physics this applies to lengths, masses, durations, intensity of electric current, and so on. In these cases, a property may be singled out as the unit, mapped to the value 1 of the scale, so that all comparable properties are represented by multiples or submultiples of the unit, and the generic value $X_{p,i}$ in X_p can be meant as k u; that is, the k th multiple of the unit u (e.g., 1.234 m). In such a ratio scale any similarity, such as

$$g(f(x)) = kf(x), \text{ with } k > 0$$

is an admissible transformation, so that a scale factor k is sufficient to transform for example values in meters into values in inches.

The following table, adapted from Stevens's paper, summarizes the main scale types and their defining conditions (where x is a property, v is a value, and g is a scale transformation).

This is the weakest, and therefore most encompassing, notion of measurement, also applicable to situations that someone would hardly accept as measurements at all (e.g., a person has performed a measurement because in their opinion this is less than that, so they labeled this as less than that). After Stevens, this position was given a more rigorous mathematical formalization by the so-called *representational theory of measurement*, in which the fact that a structure of properties can be mapped into a structure of values, through a

Table I The Main Scale Types and Their Defining Conditions

Type	Features	Invariance Condition	Admissible Transformations	Example
Nominal	Classifications	(Non)equivalence: $x_1 \neq x_2$	Injective transformations: if $v_1 \neq v_2$ then $g(v_1) \neq g(v_2)$	Blood type
Ordinal	Orderings	Order: $x_1 < x_2$	Monotonic transformations: if $v_1 < v_2$ then $g(v_1) < g(v_2)$	Hardness in Mohs scale
Interval	Numerical mappings without a “natural zero” and a “natural unit”	Ratios of differences: $(x_1 - x_2)/(x_3 - x_4) = k$	Linear transformations: $g(v) = k_1 v + k_2$, with $k_1 > 0$	Temperature in Celsius degrees
Ratio	Numerical mappings with a “natural zero” but without a “natural unit”	Ratios: $x_1/x_2 = k$	Similarity transformations: $g(v) = k v$, with $k > 0$	Mass
Absolute	Numerical mappings with a “natural zero” and a “natural unit”	Properties: x	Identity transformation: $f(x) = x$	Number of alpha decays

Source: Adapted from Stevens (1946, p. 678).

representation theorem, is considered sufficient to make it a theory of measurement: If it is a consistent representation—in this algebraic sense of consistency—then it is a measurement.

Measurability

If the long scientific, technological, philosophical, and societal development of measurement is read in light of the representational approach, a clash emerges: While consistency in representation may be acknowledged as a necessary condition of measurement, its *sufficiency* would be questioned by noting that, even if all measurements are consistent representations, not all consistent representations have been historically accepted to be measurements. Further possible constraints relate to the structure (whether quantitative or not only) and to the domain (whether physical or not only) of measurable properties. In both cases, the root is plausibly to be found in the mathematics of classical Greece.

Can a Nonquantitative Property Be Measurable?

Ratios were so fundamental in ancient Greek mathematics that *rational*, that is, *based on ratio*, was meant also as *reasonable*. And ratios were construed in terms of *measures*. Indeed, according to Euclid, in Book V of *Elements*, “a magnitude is

a part of a(nother) magnitude, the less of the greater, when it measures the greater,” whereas Aristotle, in book V of *Metaphysics*, explained that “a quantum is a plurality if it is numerable, a magnitude if it is measurable. *Plurality* means that which is divisible potentially into non-continuous parts, *magnitude* that which is divisible into continuous parts.” On this matter, distinguishing discrete versus continuous properties sounds today immaterial (hardly would someone claim that, say, the discovery that energy is quantized implies that it is countable but not measurable). Rather, what is important here is the nature of the entities whose measurability is considered. And that these concepts of magnitude and measure relate to mathematical, and not empirical, entities is clear from the fact Euclid himself considered that “a number is part of a(nother) number, the lesser of the greater, when it measures the greater.” Two millennia later, Augustus De Morgan was keen in pointing it out: “the term ‘measure’ is used [by Euclid] conversely to ‘multiple’; hence [if] A and B have a common measure [they] are said to be commensurable” (1836, p. 9).

In the past, the question whether the geometry of *Elements* is mathematical or empirical was possibly considered ill-posed, but with the break produced by the discovery of non-Euclidean geometries it became critical. And, despite some still existing stereotypes, the conclusion is clear:

Euclid was concerned about the mathematical concept of *measure*, not the empirical concept of *measurement* discussed here. Hence, the source of the *special efficacy* of measurement cannot be the Euclidean concept of measure.

Can a Nonphysical Property Be Measurable?

The emphasis on empirical entities of the experimental method promoted by Galileo Galilei led to reinterpret numerable pluralities and measurable magnitudes as physical quantities. In fact, a trace of the Euclidean focus on geometry remained in the traditional term “weights and measures,” as if weights (and pressures, temperatures, etc.) were not measurable because only dimensional quantities are. For example, in his *Mathematical and Philosophical Dictionary*, 1795, Charles Hutton called *observation* the process of evaluating a temperature by means of thermometer: It was not a measurement because only continuous geometric quantities were considered to be measurable. With such a strict implied meaning, it is not surprising that Otto Holder’s seminal paper (1901) *Die Axiome der Quantitat und die Lehre vom Mass* (translated as *The Axioms of Quantity and the Theory of Measurement*) maintained continuity and additivity as characterizing features of measurable properties. That most, if not all, psychosocial properties are not continuous, nor in fact additive, can be taken for granted: in no way is the intelligence of two individuals the sum of their intelligences, and so on (peculiarly, in those years physicists were discovering the discontinuity also of some physical quantities, at the fundamental level of quantum mechanics).

The measurability of such properties was the conundrum that led a committee appointed in 1930 by the British Association for the Advancement of Science to study “quantitative estimates of sensory events.” The conclusion was, substantially, that for a property to be measurable it must exhibit an additive structure or must depend functionally on one or more additive quantities (what Campbell had called *fundamental magnitudes* and *derived magnitudes* respectively). These conditions are possibly fulfilled by some psychophysical properties—along the line of the work started in the mid-19th century by Gustav Theodor Fechner

about the relationships between physical stimuli and the associated human responses—and in the mid-20th century triggered the development of methods to identify a hidden (interval) ratio structure in properties, particularly Rasch methods and the so-called *conjoint measurement methods*. But it also created a barrier that appeared impossible to overcome for many psychosocial properties, particularly in the context that we now call psychometrics.

A More Encompassing Concept of Measurability, Then?

The representational approach mentioned above proposed a constructive strategy to overcome these limits, at the price of neglecting (i.e., abstracting from) the way the process of measuring is structured and actually performed, a choice that makes the approach general but also generic. For sure, quantitative evaluations convey more information than non-quantitative ones, as the theory of scale types shows; but the multiplicity of the scale types also highlights the conventionality of the position according to which only quantitative properties, and possibly only properties evaluated on ratio scales, are measurable. Even in the context of fundamental physical metrology additivity and continuity are no longer taken as necessary conditions of measurability, as witnessed for example by the *International Vocabulary of Metrology*, which assumes that also ordinal properties are (quantities and are) measurable.

The clash between the strict conditions implied by the historical conception of measurement and the loose perspective of representationalism is not solved: both demarcation (to distinguish measurement from other information-related processes) and qualification (to distinguish good from bad measurements) still require a synthesis.

Toward a Theory of the Process of Measuring

Measurement is a process: hence, an actual theory of measurement, and not of measure, must be a theory of the process of measuring, and not only of the input–output relations that the process

establishes. In reference to the metaphor of the black box, a theory of measurement has to deal also with what there is inside the box, and not only with what enters and leaves the box.

As a generic basic evaluation equation, $P(o) = X_{p,i}$ in X_p , shows (a quantitative example being *length (rod a) = 1.234 m*), measurement provides information about the property P of an object o , by stating that $P(o)$ is a given $X_{p,i}$ and not any other $X_{p,j}$, $j \neq i$, in X_p . There are fundamentally two strategies to produce this kind of information. First, the value of $P(o)$ can be computed via a function of already available values (e.g., according to Newtonian mechanics the value of the length of the path traveled by a body moving at constant speed can be computed by multiplying the values of the speed and of the duration of the motion). Second, the value of $P(o)$ can be obtained by comparing the object o with other objects s_j , called *measurement standards*, each of them having a property $P(s_j) = X_{p,j}$ in X_p , so that $P(o) = X_{p,i}$ in X_p for a given i if $P(o) = P(s_i)$. For example, *length (rod a) = 1.234 m* if there is a measurement standard s_i whose length is 1.234 m and is the same as the length of *rod a*.

Within the Galilean tradition, the *special efficacy* of measurement is a matter of somehow identifying the *measurand*; that is, the property intended to be measured, and finding its value in the given scale in the here-and-now conditions of the process. To this goal, some calculations may be useful or even required—for example, to correct known environmental effects without an empirical intervention—but the first strategy is not sufficient: a pen-and-paper-only evaluation is not a measurement, because a measurement requires the property-related direct or indirect empirical comparison of the object under measurement and the measurement standards that materialize the chosen scale. (In the foregoing example, in order to consider that the length of the path is measured—in what it is traditionally called an *indirect measurement*—and not just computed, the information about the speed and duration of the motion must be acquired in the actual conditions, something that a computation cannot provide).

The two key features to qualify a measurement derive from this. First, measurement results are expected to be *object related*, by reporting the

outcome of the comparison of the object under measurement and some measurement standards. This explains the importance attributed to constructing and characterizing the behavior of measuring instruments, which have the goal of guaranteeing a sufficient object relatedness by removing the effects of all influence properties. Second, in most real-life cases measurements are expected to produce information that is socially transferable, and therefore interpretable in the same way by different individuals in different times and places, a condition that makes it *subject independent*. This explains the importance attributed to disseminating measurement standards and calibrating measuring instruments against them, so that the measurement results are metrologically traceable to the chosen scale, and in particular to the definition of the unit in the case of ratio quantities.

The quality of measurements and their results is fundamentally assessed about their object relatedness and subject independence, *objectivity* and *intersubjectivity* for short, through the several (usually quantitative) properties that characterize the process in its multiple facets: accuracy, specified as precision and trueness, in turn specified as sensitivity, selectivity, stability, resolution, and so on. The overall summary—traditionally related to measurement error—has been focused on *measurement uncertainty* in the several decades, hence with an encompassing epistemic emphasis on the information produced by the process and its trustworthiness: Measurement is thus conceived of as an empirical and informational process that is designed on purpose, whose input is an empirical property of an object and which produces values of that property that are explicitly justifiable in their (good or bad) objectivity and intersubjectivity.

Luca Mari

See also Accuracy; Applicability of Mathematics in Science; Data and Phenomena; Epistemology; Information Theory; Modeling; Scientific Realism; Statistics

Further Readings

- Campbell, N. R. (1920). *Physics: The elements*. Cambridge University Press.
- Joint Committee for Guides in Metrology. (2008). *Evaluation of measurement data—Guide to the expression of uncertainty in measurement (GUM)*.

https://www.bipm.org/documents/20126/2071204/JCGM_100_2008_E.pdf/cb0ef43f-baa5-11cf-3f85-4dcd86f77bd6

- De Morgan, A. (1836). *On the study and difficulty of mathematics*. Library of Useful Knowledge.
- Joint Committee for Guides in Metrology. (2012). *International vocabulary of metrology—Basic and general concepts and associated terms* (VIM; 3rd ed.; 2008 version with minor corrections). https://www.bipm.org/documents/20126/2071204/JCGM_200_2012.pdf
- Kuhn, T. S. (1961). The function of measurement in modern physical science. *Isis*, 52(2), 161–193.
- Kula, W. (1986). *Measures and men*. Princeton University Press.
- Mari, L., Wilson, M., & Maul, A. (2021). *Measurement across the sciences—Developing a shared concept system for measurement*. Springer.
- Michell, J. (1999). *Measurement in psychology: A critical history of a methodological concept*. Cambridge University Press.
- Rossi, G. B. (2014). *Measurement and probability—A probabilistic theory of measurement with applications*. Springer.
- Stevens, S. S. (1946). On the theory of scales of measurement. *Science*, 103(2684), 677–680.
- Tal, E. (2020). Measurement in science. In E. N. Zalta (Ed.), *The Stanford encyclopedia of philosophy*. plato.stanford.edu/archives/fall2020/entries/measurement-science

MEDICINE, ANTIQUITY

Throughout the ancient world, many different theories were devised concerning the nature of illness, injury, and death. This resulted in the development of a wide range of methods of treating afflictions. Although many ancient societies also believed in some form of afterlife, the solace this provided could only go so far toward alleviating the grief and mourning of families and friends upon the death of an individual. This entry describes and explains the concepts and practice of medicine in the ancient civilizations of Mesopotamia, Egypt, Greece, Rome, China, and India, noting their influences on subsequent eras down to the present day.

Discoveries of fossilized teeth from prehistoric times show problems of erosion and tooth decay, with bone loss attributed to poor diet. The oldest

surviving item that shows some knowledge of medicine is a statue found in *Çatalhöyük* in present-day Turkey, dating from ca. 6500–5700 B.C.E. and depicting a seated female figure. It is thought to be a representation of the goddess of fertility giving birth and shows the mother seated on a specially designed chair.

Mesopotamia

Not much is known about medical treatments in ancient Mesopotamian empires of Sumer, Babylonia, and Assyria. There is one surviving piece of Akkadian literature, however, that focuses on the story of Gimil-Ninurta, which involved heating water over a fire. In the story, however, this is actually used as a ruse to rob a town mayor rather than to cure him of illness. Illnesses were commonly regarded as punishments from the gods, and as a result, there were many deities to which offerings would be given in the expectation of being rewarded with good health and prosperity. One Assyrian bronze amulet, now held by the Louvre Museum in Paris, shows the exorcism of a demon. In the Babylonian Code of Hammurabi (ca. 1772 B.C.E.), a legal document carved into a massive stone pillar, rules governing medicine are scant and refer chiefly to the outcome of operations, not to the medical and surgical procedures themselves. The code does mention removal of tumors, which had to be treated with a metal knife, and also mentions diseased flesh on a broken bone, possibly a reference to gangrene. From this and other information, however, it can be seen that there were three types of priests, one of which dealt specifically with the sick, having been taught at schools connected to temples. The Code of Hammurabi lays down penalties for medical practitioners if they failed to achieve a cure. These range from fines to amputation of hands if a patient died in surgery or lost his eye as a result of an operation. There is also mention in Babylonian texts of veterinary doctors who were regarded as having lower social status than physicians who treated humans.

Egypt

There were specialist physicians in ancient Egypt, even during the Old Kingdom (ca. 2700–2200

B.C.E.), and certainly there were doctors assigned to the rulers and who were known as “chief physician to the Pharaoh” and by similar titles. The demigod called *Imhotep* was believed to have been the creator of medical knowledge. Nevertheless, and in spite of the great skills that the Egyptians developed in embalming the dead, there were evidently severe limits to their anatomical knowledge. They obviously knew about the bones, as well as the larger organs such as the stomach, heart, and spleen. Some ancient Egyptian records do state that the veins and arteries—which were known to be linked to the heart—were believed to carry water, air, and also excretions through the body. Their anatomical knowledge, however, seems to have ended there.

Essentially, the major theories concerned with the treatment of medical afflictions in ancient Egypt rested on the *members* (organs) which were connected to the heart. One ancient text refers to 12 members being in pairs located in, and regulating, the breasts, legs, forehead, and other parts of the body. A later text mentions some 40 of these. The theory was that these were the root of many medical problems when they became blocked, heated, or extremely stiff. Thus, the medicine or treatment prescribed was intended to deal with these members.

Gradually, physicians in ancient Egypt came to believe that they could diagnose patients by seeing them without any detailed examination. This led to the compiling of medical texts that prescribed particular treatments. One of these notes, “When thou findest a man who has a swelling in his neck; and who suffers in both his shoulder-blades as well as in his head, and the backbone of his neck is stiff, and his neck is stiff so that he is not able to look down upon his belly,” he is to be told the importance of the swelling on his neck and ointment should be rubbed on it, after which he should be healed. Many of the medicines used by the Egyptians seem to have come from the olive tree, but there were also many other preparations that came from berries and other fruits, as well as herbs, with very few made from animal ingredients.

Greece and Rome

The ancient Greeks gained much of their initial knowledge of medicine from the Egyptians but

developed this considerably. The Greek physician Hippocrates of Cos (ca. 460–ca. 370 B.C.E.) is regarded as the *Father of Medicine* because his teachings and theories changed the view of many Greeks and established medicine as a separate discipline, distinct from philosophy. The first aspect of treating a patient was, according to Hippocrates, to observe and examine the patient in detail before making a diagnosis. Many of his cures depended on rest, exercise, and a proper diet, although he also identified a number of herbs and plants from which medicines for treating particular ailments could be acquired. Hippocrates also believed that it was important to keep detailed notes on the treatments, which would allow doctors to remember what had previously been prescribed for a particular person, and also to gain more knowledge about the success or failure of individual cures. Hippocrates wrote down many of his ideas and cures, and these came to be known as the *Corpus Hippocraticum*, or Hippocratic collection. He is also credited with the Hippocratic Oath, still in use today, by which physicians must promise to uphold specific ethical standards during their treatment of people.

From the 5th century B.C.E., the influence of the Greeks spread throughout the Mediterranean and then through the Persian Empire. By around 450 B.C.E., the Greeks had established Asclepios as the god of medicine and healing. He was claimed to have been the son of the god Apollo and a mortal mother, but some scholars believe that there might have been an actual person called Asclepios, and indeed there is some allusion to this in the works of Aristophanes (ca. 446–ca. 386 B.C.E.). Either way, buildings or temples known as Asclepeia came to be built in many towns and cities where people came to be cured. The curing of patients was seen as being partly the work of the doctors at the Asclepeia, and partly through divine intervention of the god to whom they brought gifts of food, which were placed on the altar. One interesting development was that patients themselves were encouraged to eat well and thus build themselves up physically, as well as to wash themselves in a bath house attached to the temple. Gradually, these practices became associated with two goddesses, Hygeia and Panacea, who are claimed to have been the

daughters of Asclepios and whose names are preserved in the terms *hygiene* and *panacea*. The Greek theory that there were both natural cures and supernatural cures continued through Roman times and remains to the present day with, for example, the Christian concept of miraculous cures.

Although medical knowledge at that time was capable of dealing with straightforward physical ailments such as broken limbs, cuts, and abrasions, there was a growing body of knowledge which showed that there were also mental ailments, which often exacerbated standard medical ones. Hippocrates came up with the theory that the body consisted of four *humors*—black bile, yellow bile, phlegm, and blood—although this later became associated with the Greco-Roman writer and physician Galen. According to the theory, people would remain healthy when the four humors were in balance. Yellow bile, representing fire (or summer), made people irritable or angry. By contrast, black bile, representing earth (or autumn), made people sad or melancholic. Phlegm, representing water (or winter), made people sluggish; blood, representing air (or spring), made people more optimistic.

On this view, if the humors were out of balance, then so was the mental and physical state of an individual. It was therefore important for doctors to try to achieve the proper balance. Occasionally, there were natural ways such as vomiting, and also the coughing of phlegm, sneezing, and nosebleeds, which were all attributed to the body itself trying to solve problems. Hippocrates mentions the need to drain the lungs of patients who are suffering from pneumonia. The theory of humors was to remain central to medical beliefs in Europe until the Renaissance. It was clear that more people tended to get colds during the winter, and some suffered from allergies to pollen (*hay fever*) during spring, with all these ailments accounted for by the belief in the importance of balancing the humors. There were obvious problems with this, such as during an outbreak of the plague in Athens in 430–427 B.C.E., when many people died and it was widely doubted that so many Athenians could have had their humors out of balance at the same time. This led to ideas that either the Spartans had poisoned the wells or that the gods had turned

against Athens, with others suggesting that poor air quality was responsible. Indeed, this idea of bad air pervades through ancient Europe, with people living near swamps tending to be more affected than those living in drier areas. This was to give rise to those suffering from *bad air* developing what became known as *malaria*. It was not until the 19th century that this came to be connected with the mosquitoes that flourished in swampy regions.

Archaeologists have found surgical instruments from the ancient world, and combined with surviving texts, it is clear that surgery was seen as important to the Greeks. Scholars have calculated that Homer, in *The Iliad*, mentions some 147 battle wounds in his text, with 106 of these being made by thrusts from swords, 17 from slashes by swords, 12 from piercing by arrows, and the remaining 12 by projectiles from slingshots. Most of these who were wounded subsequently died of these wounds, but an exception was King Menelaus, who was hit by an arrow in his waist. There is a reference to the doctor removing the arrow and then cleaning the wound, after which a healing ointment is applied. Similarly, there are references in later texts to limbs being removed after gangrene had set in.

Many Roman surgical tools have been found in archaeological *digs*. With their large armies, the Romans had significant numbers of army doctors for treating the many injured soldiers. However, little or nothing was known about germs and infections, with the result that surgical tools for treating the rich were often elaborate and had inlaid patterns that harbored plentiful microorganism, whereas simple tools used on poorer people were often cleaner.

Some medical theories were advanced by non-doctors. The dissection of human bodies was illegal in Greco-Roman times, and indeed this remained the case throughout medieval Europe. As a result, many doctors were clearly ignorant of the exact nature of the human body. This led the philosopher Aristotle (384–322 B.C.E.) to suggest that doctors should dissect animals to get a greater understanding of the human body—and it was Aristotle who is attributed with having come up with the concept that the heart was the central organ in the body and that the heartbeat was the sign of life. Gradually, there were more advances

in medical theories of the Greeks. Herophilus (335–280 B.C.E.) did dissect some human bodies and concluded that the brain and not the heart controlled the body, although the concept of the heart being the focus remained for centuries, especially for romantic attachments and in poetry. Erasistratus (304–250 B.C.E.) examined the heart and found valves but was unsure whether the heart pumped air or blood.

During the Roman Republic and the Roman Empire, there was a growth in cities, and one of the major problems that emerged was epidemics, especially the plague. The Romans had used their building skills to bring fresh water to cities and also to build sewers and public lavatories. They also had their own god of health, known as Salus. However, as noted earlier, in 295 B.C.E., a plague hit the city of Rome, and this led to the building of a temple in Rome dedicated to Asclepius. Even before Greece became a part of the Roman Empire, the Romans had absorbed many medical theories from the Greeks. Indeed, Greek doctors were favored during Roman times and it seems likely that some of the earliest ones were captured in war. Others moved to Rome, and indeed, Pliny the Elder (d. 79 C.E.) mentions that some Roman doctors had started taking Greek names. The old traditions relate that St. Luke was a doctor, although modern scholars disagree with this association. Claudius Galen was also Greek, having been born in Pergamon, a Greek city in Ionia (modern-day Turkey). He moved to Rome at the age of 20 and helped treat Marcus Aurelius and the imperial family during an outbreak of the plague. His work emphasized the ideas of the humors propounded by Hippocrates and others and indeed advised that doctors in training should go to Alexandria to consult the library there and also to dissect human bodies, which was evidently allowed in that city.

Women also came to be regarded as good at healing people, with Pliny the Elder writing in ca. 75 C.E. that “the science of charms and herbs is the one outstanding skill of women.” Pliny also suggests a number of cures with egg yolk mixed with wine and poppy juice being taken to treat people with dysentery and unwashed wool coated in honey being used to heal old sores. However, the theory of humors was still important; Pliny

believed that epileptics having a fit were trying to throw off an infection so as to bring the humors into balance again.

The Greeks and the Romans were also especially conscious of poisons, which in small doses could be used as a cure for disease but in large doses could kill. The Athenian philosopher Socrates was found guilty of an alleged offense and sentenced to death by swallowing a lethal dose of hemlock. Mithridates VI, King of Pontus, was so paranoid about being poisoned that he took small daily doses of certain poisons to build up his own resistance to them. Also, prisoners who were about to be executed could be used to test poisons and antidotes. This led to the development of what became known as *mithridatium*, an antidote that the king kept secret but was eventually uncovered by the Romans when they invaded Pontus. The emperor Claudius is reputed to have been poisoned with a recipe that included mushrooms, and Pliny the Elder famously mentions that there are some 7,000 different types of poisons.

China

In ancient China, there was a large range of medical procedures that were followed for centuries. Inscriptions from the Shang dynasty (1600–1100 B.C.E.) ascribe medical ailments to curses on ancestors, which are then passed down to later generations. This was a generalization based on the recognition of hereditary illnesses rather than the incidence of minor afflictions and injuries. The oldest surviving medical text in China was by Fu Hsi (ca. 2900 B.C.E.) who is believed to have devised the binary system of Yin and Yang, the fundamental, abstract, and complementary principles of the universe. When the Emperor Qin Shi Huangdi ordered the burning of books (and the slaughter of scholars) in 213–210 B.C.E., he specially excluded medical works from the destruction.

Chinese medicine had come to focus on the treatment of illnesses using herbs. Later came the development of acupuncture; massage, exercise, and dietary therapy were also important. These practices were intended to promote balance in the body between the forces of Yin and Yang, as outlined by Fu Hsi. Yang was commonly

understood as facing the sun, with Yin in the shade, or as manifested in the difference between fire and water. In many ways, the principle is not that dissimilar from the Greek views of the humors; what was essential was to attain the proper balance between these two forces. There was also the belief that those suffering from disease had tried to move from their assigned place in society, as the affliction was a reflection of a patient's context and surroundings as much as anything else.

Diagnosis was traditionally made by feeling the pulse in the right wrist and then the left wrist of a patient in the course of an examination of the skin. However, male doctors were not allowed to examine a naked woman; rather, female patients pointed to their symptoms and affected bodily areas on dolls that doctors kept for that purpose. And the Chinese easily differentiated between the epidemics such as the regular occurrences of the plague, the seasonal illnesses, and those which involved treating symptoms over long periods, especially for mental and nervous disorders.

The Chinese of ancient times had discovered much more about the circulation of the blood than was the case in Europe. The Chinese classical medical text, the *Nei Ching*, from around the 2nd century B.C.E., mentions that the blood flows through the body in a *continuous circle*. There was also a belief that some vessels carried air, although the concept of osmosis, and also that the blood-carrying vessels were closed, was not yet recognized.

The Chinese practiced acupuncture from the time of the Han dynasty. It is said to have its origins in some soldiers in battle being wounded with arrows but having been cured of other illnesses, and as a result, considerable research was undertaken in this field. Some pictographs from the Shang dynasty and also early surviving needles are likely to be connected with blood-letting and lancing wounds rather than acupuncture, which was originally done with bone needles from at least 200 B.C.E. The theory behind acupuncture is that, because the flow of *chi* (vital energy) is thereby altered, specific parts of the body are stimulated by insertion of the needles, although the respective insertion sites may be

remote from the affected organ or region. This therapy has been found to be effective in reducing some levels of pain, and it has remained popular since ancient times as an alternative method of medical treatment. Traditional Chinese medicine using herbs and other preparations has also continued to prove to be effective in dealing with certain conditions.

India

In the south Asian region of India, medical beliefs focused on what the ancient Indians saw as the five classical elements: earth, water, fire, air, and ether. They also recognized seven basic tissues: plasma, blood, muscle, fat, bone, marrow, and semen. The aim was to remain healthy by keeping a balance between these through moderate behavior. In many ways, this was similar to classical Greco-Roman and Chinese views. Indian views on afflictions were that some arose from a lack of moderation in connection with rest, diet, or other factors such as sexual intercourse. They also felt that some natural urges—such as to sneeze—should not be suppressed but rather indulged as this was a natural way in which the body tries to achieve harmony. Massage and rest were particularly important cures, with a large range of medicines made from parts of plants such as bark, roots, or leaves, as well as from fruit and seeds.

Ancient Hindus had a pantheon of gods, with Agni emerging as the Hindu god of fire, and who was believed to be able to cure fevers. And the earliest complex medical equipment in the world was found by archaeologists who discovered a dental drill dating to as long ago as 7000 B.C.E. used by the Indus Valley civilization in modern-day Pakistan. It consisted of a hand-operated drill which has flint points and can be used to remove decaying dental tissues. Indian medicine also placed particularly heavy emphasis on hygiene, including regular bathing and also the daily cleaning of teeth.

Justin Corfield

See also Medicine, 19th Century; Medicine, Contemporary; Medicine, Enlightenment; Medicine, Middle Ages; Medicine, Renaissance

Further Readings

- Cruse, A. (2004). *Roman medicine*. Tempus.
- Dawson, I. (2005). *Greek and Roman medicine*. Enchanted Lion Books.
- Elliott, J. (1971). *Outlines of Greek and Roman medicine*. Milford House.
- Erman, A. (1971). *Life in ancient Egypt*. Dover Publications.
- Gabriel, R. A. (2012). *Man and wound in the ancient world: A history of military medicine from Sumer to the fall of Constantinople*. Potomac Books.
- Huard, P., & Wong, M. (1968). *Chinese medicine*. McGraw Hill.
- King, H. (2001). *Greek and Roman medicine*. Bristol Classical Press.
- Kutumbian, P. (1999). *Ancient Indian medicine*. Orient Longman.
- Loewe, M. (1968). *Everyday life in early imperial China during the Han dynasty, 202 B.C.E.–C.E. 220*. B. T. Batsford.
- Mattern, S. P. (2013). *The prince of medicine: Galen in the Roman Empire*. Oxford University Press.
- Moscatti, S. (1968). *The world of the Phoenicians*. Weidenfeld & Nicolson.
- Paoli, U. E. (1963). *Rome: Its people, life and customs*. Longmans.
- Phillips, E. D. (1973). *Greek medicine*. Thames & Hudson.
- Saggs, H. W. F. (1967). *Everyday life in Babylonia and Assyria*. B. T. Batsford.
- Scarborough, J. (1969). *Roman medicine*. Cornell University Press.

MEDICINE, CONTEMPORARY

Medicine is usually defined as the practice concerned with the preservation of health and the prevention, alleviation, or cure of disease. Contemporary medicine is also known as *allopathic* medicine. The term comes from the ancient Greek prefix ἄλλος, *other*, *opposite*, and the suffix πάθος, *to suffer*. In short, it refers to treating a symptom with its opposite as, for instance, treating an infection with antibiotics or curing constipation with a laxative.

The term *allopathic* has long been considered derogatory by many in mainstream medicine, as much of standard medical practice does not seem to work in that way. Nonetheless, one can say that contemporary medicine is focused on curing patients, intervening by *removing* the cause of

the disease. If that is the main goal of medicine, then it does require an understanding of basic concepts such as *health* and *disease* to identify who needs help, and more complex ones, such as *cause* and *evidence*, to know how to intervene efficiently. Theories of health and disease, and theories of evidence and medical knowledge, have profound societal implications because whether and how an individual has to be treated depends mostly on their content. While it is certainly true that medicine is a practical and societal enterprise, it is at the same time a scientifically founded discipline. Furthermore, in medicine one can find theories in both a foundational and a special sense. Foundational theories lay the groundwork of the discipline, while special theories aim at providing an account of observed phenomena. Notably, both the categories of medical theories are strictly connected with practical action (understanding, treating, and preventing disease), unlike most other scientific disciplines wherein theory is more distant from applications (consider, e.g., theoretical physics). In fact, the theory of medicine is subordinate to the constitutive ethical imperative of medicine, which demands action to help people in case of sickness or accident. Nowadays, no other scientific discipline has an influence on ordinary life even remotely equal to that of medicine.

Theories of Health and Diseases

Medicine serves a very practical purpose: to heal the sick. But to accomplish this goal, a theory to distinguish sick from healthy individuals is needed. That has indeed been one of the main theoretical goals of medicine, since the humoral theory associated with the ancient Greek physician Hippocrates (ca. 460–375 B.C.E.) and the eclectic theory of the best constitution associated with the physician and philosopher Galen (129–216 C.E.), in the time of the Roman Empire. For many people, the disease is a given, a state of affair easy to identify. That is not actually the case, though.

In the last two centuries, there has been a strong medicalization of society. Some conditions that were previously not considered illnesses became recognized as medical conditions. Today, medicine is prevalent in nearly every area of our lives.

Medicalization, taken at face value, is not inherently a negative idea, and it has certainly enhanced our quality of living. But the basis on which human problems are to be described and handled as medical issues must be very clear. Otherwise, vested interests can distort the benign essence of medicine. Indeed, the definition of the terms *health* and *diseases* has been a central topic of theoretical reflections on medicine since the 1960s.

There are three major competing theories of health and diseases in contemporary medicine: naturalism, constructivism (or normativism), and hybridism. Some scholars, often called *naturalists*, argued that definitions of health and diseases must be based on biological facts. Only science can tell what counts as a disease, and the more we know about human biology the better the definition. According to *naturalism theory*, health and disease are empirical, objective, and value-free concepts, and thus they are to be understood in purely scientific terms. The most prominent naturalist account of health and disease is perhaps Christopher Boorse's *biostatistical theory*, developed in the late 1970s. Within this account, a person is healthy when she has normal biological functions. And normality is defined by observed frequencies in a specified reference class (e.g., within a specific age group and sex). The theory needs appropriate reference classes because human beings have wide variability in biological functions: what is normal in one group can be abnormal in another. For instance, normal levels of estrogens in adult females are different from normal levels in infant females. A young woman who has the same level of estrogens that is normal for an older woman would be deemed diseased. So, to assess when a process or a part of the organism, such as an organ, functions in a normal way, one must compare it with the organs of similar people in the same age group. If it works similarly, then the organ is normal, and the individual is healthy; otherwise, the organ is diseased. Indeed, according to the biostatistical theory, disease is just a departure from health or normality.

The main alternative approach to the naturalist theory of health and diseases is *constructivism* or *normativism*. Most proponents of this theory agree that the terms *health* and *disease* must be defined explicitly and that definitions strongly depend on values. On this theory, being healthy

requires the possession of various capacities, not only normal biological functions. This theory is epitomized by the World Health Organization (WHO, 1948) definition of health: "a state of complete physical, mental and social well-being and not merely the absence of disease or infirmity." This definition expands the concept of health beyond biological considerations. The concept of health seems then closer to the idea of well-being, which is an evaluative notion. At the same time, the concept of diseases is not confined only to biological dysfunctions. On the extreme formulation of the theory, the *constructivists* believe that the definition of disease is nothing but a result of our value judgments. More specifically, the diseases are those conditions we as a society do not like much. Therefore, this account seems better to grasp all those diseases that do not have a clear biological basis—such as depression for instance. Further, it can accommodate individuals' feelings and culture-bound syndromes. However, unlike naturalism, the constructivist theory is not a unified account and there are many possible varieties, which cannot be discussed here owing to space limitations.

Finally, the *hybrid theories* of health and disease combine some aspects of both theories, in the attempt to overcome some of their major issues. Since both of those accounts suffer from serious objections, some scholars (e.g., Jerome Wakefield) have developed hybrid theories taking important elements from both naturalism and constructivism. From naturalism, hybridism takes the idea that health and disease must involve the biological functioning of the organism. From constructivism, hybridism holds that biological dysfunctions must be disvalued for a condition to be considered a disease. On this account, biological abnormality, although important, is not sufficient for a condition to be a disease, nor is it sufficient for there to be the mere capacity of a condition to cause harm. Therefore, the hybrid theory holds that there are two necessary conditions that a state must have so as to be deemed a disease: It must be a biological dysfunction and the state must be disvalued or harmful, but neither one is sufficient. Even though hybridism inherits some of the drawbacks of naturalism and constructivism, it has an additional advantage: It is more consistent with the actual medical practice.

All in all, however, these efforts to establish theories of health and diseases have demonstrated that there is no common concept that best encompasses their essence. New interpretations are introduced as either science or behaviors change over time, giving to these terms an irreducible complexity that demands continuous analytical attention.

Theories of Causation

If the central goal of medicine is to help the sick, then it is fundamental to determine what knowledge is required to achieve this goal. That is why all the medical theories in contemporary medicine are theories that aim to explain or understand a disease in order to treat it. In contemporary medicine, explaining or understanding a disease is all about recognizing its cause. But to develop such a piece of knowledge, it is essential to know some basics about the general nature of causation.

There are three main theories of causation in medicine. One is the *sufficiency–necessary theory*, which holds that a cause is a condition either necessary or sufficient (or both) for its effect. Necessary means that the disease Y will never develop without the condition X. Sufficient means that every time the condition X is present, the disease Y will always develop. Three different types of causal relations can then be derived: necessary and sufficient, necessary but not sufficient, sufficient but not necessary. In all the possible scenarios we can then apply this causal theory to identify a condition as the cause of a particular disease. This view is traditionally associated with the germ theory of disease, wherein diseases are caused by specific agents. This very important theory in the history of Western medicine is attributed to a German physician and microbiologist of the 18th century, Robert Koch, who developed the famous four postulates of disease that link a particular disease to a very specific pathogenic microorganism. However, physicians have understood, at least since the advent of the age of chronic diseases, that not all diseases can be attributed to a particular, necessary, or sufficient cause. Cigarette smoke, for example, is not necessary for the occurrence of lung cancer, nor is cigarette smoke its sole cause. In fact, the causes of complex chronic diseases, such as cancer and heart disease, appear to fall under the “neither necessary nor sufficient”

category. But even in infectious diseases, the presence of a pathogenic agent is not a necessary nor sufficient condition, since in some context the pathogen alone is not able to cause the disease, as for instance in asymptomatic carriers, people who have actually been infected but do not display the symptoms of the disease.

The second theory is the so-called *probabilistic theory of causation*. On this account, a cause is a condition that raises the probability of developing the disease. A cause is, therefore, neither necessary nor sufficient, but it just increases the probability that its effect will occur. For example, smoking is not necessary nor sufficient for developing lung cancer: Some smokers never develop it, and some nonsmokers develop the disease nonetheless. But smoking certainly increases the probability of the occurrence of the disease. Within this theory, the concept of cause goes along with the idea of risk factor. Epidemiologists use this term to indicate a condition that is probabilistically associated with a given outcome. But not all risk factors are necessarily also causing the disease in the strict sense (e.g., low socioeconomic status is often associated with an increased risk of cancer, but it is not a cause of the disease). While this theory might sound appealing, it might not always be able to distinguish a genuine relationship from spurious correlations. For example, many observational studies have found a correlation between drinking alcohol and cancer. Cancer could be due, in fact, to drinking, but it could be also due to some features of drinkers, such as poor diet or smoking. Answering the question of whether a given factor is or is not a cause requires making a judgment. There are no rigid and objective criteria for determining whether a causal relationship exists; nonetheless in 1965 Austin Bradford-Hill proposed a list of nine aspects of associations that should be measured when attempting to draw conclusions regarding causality, which has since become very popular among practitioners. The list includes the following nine aspects: the strength of association, consistency, specificity, temporality, biological gradient, plausibility, coherence, experiment, and analogy.

The third theory is the *sufficient-component theory of causation*. According to this view, the cause of a disease is the total collection of

different conditions that are together sufficient for the occurrence of the disease. So, a sufficient-component cause is like a pie, composed of several different ingredients, none of which is sufficient for the disease on its own. However, when all the ingredients are present, the disease will occur. Each of the ingredients of the pie is nonetheless necessary on its own to the cause, since in fact by removing one of them the disease will not occur. This view has been promoted by an epidemiologist, Kenneth Rothman, based on a philosophical account of causality developed by John Leslie Mackie. According to Mackie, a cause is an insufficient but necessary part of an unnecessary but sufficient condition for an effect. That is why a cause is not necessary *per se*, since more than one set of components may be sufficient for the same effect; that is, a disease may have multiple causes. Considering once again the smoking–lung cancer example, smoking has to be combined with other necessary components (e.g., genetic predisposition) in order to cause lung cancer.

Recognizing genuine causal relationships from spurious associations is of course paramount to the success of medical practice, as treatments and public health policies are addressed mostly to remove the causes of diseases. So, theoretical reflections on causality in medicine are not only interesting *per se* but may also impact on the well-being of patients.

Theories of Medical Knowledge and Evidence

Everyone agrees that medical doctors should act in accordance with scientific evidence but disagrees about what evidence precisely is. Given a medical problem, what is the best way to solve it? What does count exactly as scientific evidence? Answers to this kind of questions are usually a matter for philosophers and epistemologists. Yet some medical practitioners at the beginning of the 1990s took these issues with regard to medicine very seriously as well. The *evidence-based medicine* (EBM) movement was the result of that reflection, and it has revolutionized and shaped contemporary medical practice. Its proponents have established a new medical epistemology according to which medical decisions must be

made on the basis of a very particular type of evidence, more precisely that provided by randomized controlled trials (RCTs).

One of the most important types of research in medicine is indeed to find effective ways to intervene in the disease. Therefore, the problem for medicine is what knowledge is valid to assess the effectiveness of interventions. For hypotheses regarding the effectiveness of interventions, there is indeed a great variety of evidence, sometimes even conflicting with each other—from *in vitro* and *in vivo* research, to observational studies, to case reports and individual experience of physicians. However, not all kinds of evidence are equally valid. In fact, according to EBM, medical research is prone to bias much more than any other scientific enterprise. Thus, evidence should be ranked according to its potential to prevent bias, and this potential depends eventually on the research design. Study designs that rank higher on the hierarchy of evidence are more likely to provide reliable evidence. Over the years, stemming from the original work of David Sackett (the physician considered to be the father of EBM), researchers have identified a list of biases that can occur when experimenting and have agreed on a list of debiasing controls and procedures to be implemented. From an ideal point of view, the methodology of RCTs can deliver more unbiased results, at least when compared with other study designs. This is because there is more room to implement debiasing procedures in RCTs than in any other study.

The theory of medical knowledge and evidence-based contemporary medicine rank scientific study types in a very rigid hierarchy. On top of the so-called *evidence pyramid*, there are meta-analyses and systematic reviews of RCTs (a way to summarize research results). Single RCTs are just below, followed by cohort studies, case-control studies, case series, and case reports. Surprisingly, mechanistic studies (i.e., laboratory basic research conducted *in vitro* and *in vivo*) lie instead at the bottom, because they are considered less relevant for clinical research than any other kind of study, even if they are fundamental for understanding diseases.

Despite its appeal, EBM continues to invoke many criticisms. According to many scholars, the EBM theory of evidence is too narrow as it excludes information important to medical

practitioners, such as professional experience and patient-specific factors. For this reason, the usefulness of EBM application to individual patients is considered limited. And because there are so many uncommon diseases and variants, for many patients we do not have any evidence, since actual cases hardly match with the patients included in experiments. Nonetheless the criticisms, EBM is still the dominant paradigm in the field, more than 20 years after its conception.

In the end, EBM makes a normative argument as to whether it is truly possible to take certain kinds of evidence as reliable medical knowledge. In this sense, it is a theory of knowledge. But it also takes a normative stance on medical practice: Wherever possible, medical decisions should be based on the best available evidence (i.e., RCTs) about the available interventions. EBM places medical knowledge at the center of medical practice, adopting a very strict theory of what can be regarded as medical knowledge.

This appeal for a *theory of medical knowledge and evidence* is very important to patients. Failure to determine causal arguments can indeed be harmful—contributing, for example, to the administration of ineffective treatments. Also, medical errors are especially troubling for medicine itself, because they are *visible* to the public, and too many mistakes will lead to the growth of medical skepticism in society. Having a clear theory of medical knowledge and evidence will not only help doctors make better choices but also strengthen medical practice.

Reductionism and the Biomedical Model

Along with the broad theories composing its foundations, contemporary medicine has also produced many narrower special theories aiming at understanding and explaining phenomena of medical interest. Among the latter, for instance, neurophysiology theory explains how biochemical transmitters may affect mood, behavior, and pain; mechanical laws prescribe the functioning of muscular activity; immunological theory explains how vaccines work; theories of cancer metabolism elucidate tumor growth; and so on. Such theories derive from contemporary laboratory life sciences, such as molecular biology, which are conducted with the methods of the exact sciences and the vital support of new

research technologies. This is because contemporary medicine widely embraces *reductionism*.

In a nutshell, reductionism is the philosophical idea that all higher level (e.g., social, mental, and even medical) phenomena and processes can in principle be explained at a lower level. This theory is a very powerful tool for investigating phenomena in medicine. The basic idea of medical reductionism is that diseases are microphysiological problems to be solved by a direct intervention in the underlying biological mechanisms. Medical theories deconstruct a disease into biological parts and processes and compare them with the corresponding parts and processes of people without the disease. Accordingly, diseases are to be investigated by identifying the pathological microphysiological parts and processes that constitute them, and consequently, medical treatments should target those microphysiological parts and processes.

Reductionism in medicine is also often referred to as the *biomedical model* of medicine, because of its insistence on the biological basis of medical phenomena. This view has been the dominant trend in medical research during the last 50 years, at least since the first advances and successes of molecular biology. Many disorders have indeed been adequately explained in reductionist terms, and this in turn has led to some very successful medical treatments. For example, antibiotics, insulin, vaccines, and immunotherapies are just a few of the many success stories of reductionism.

However, despite its early undeniable successes, the biomedical model has been widely criticized. Its opponents argue that there has been an overemphasis on the pursuit of genetic or molecular-level explanations of disease to the exclusion of alternative, yet significant, levels of explanation. Moreover, according to its critics, the biomedical model leaves out important information about the disease at the phenotypic level that is essential to the pursuit of effective treatments. For instance, many diseases that have been viewed as *genetic*, such as cystic fibrosis and Alzheimer disease, cannot actually be reduced to unified entities with a specific genetic cause. Moreover, since there is probably a genetic variation component in susceptibility to virtually all diseases, the very concept of genetic diseases dissolves. Furthermore, according to the critics,

reducing the differences between individuals merely to their respective genomes is impossible, since all phenotypic traits are the result of an interaction between genes and the environment within which they are expressed. From this perspective, a more *holistic* approach that takes into account social and environmental factors can deliver a better understanding of diseases, as well as better treatments.

Criticisms notwithstanding, today the reductionist theory and the biomedical paradigm remain the dominant approach to medical research and practice. In the early 21st century, there has been even greater emphasis on reducing diseases to the genetic and molecular level. In fact, the biomedical paradigm has become even more popular, mostly due to a series of significant achievements in biomedical research, such as the completion of the *Human Genome Project in 2003*, identifying and mapping all of the genes comprising the human genome.

Recently, new advances in genetics and the growing availability of health data have triggered a major improvement of the biomedical model: *personalized* and *precision* medicine. The basic idea, epitomized by the slogan “treatment for the right patient at the right time,” is that treatments can be *tailored* based on individual genetic differences, or other kinds of molecular information. This data-driven approach to treatment has been proved to be efficient, especially in oncology. One reason is that cancer is undoubtedly linked to genetic factors: Cancer harbors many mutated genes that contribute to the genesis and development of the disease. Moreover, recent technological breakthroughs, namely next-generation sequencing and single-cell sequencing, have made it possible to observe genetic variations at the single-nucleotide level between different types of tumors, as well as between individual tumors. Tumors harboring certain mutations are more vulnerable and respond better to specific therapeutic agents. And this variability can then be exploited in order to provide the best therapeutic strategy for each patient. With regard to this, the drug *vemurafenib* is a classic example, as it has been developed precisely in order to treat a very specific subset of cancer patients, those with late-stage melanoma harboring the V600 mutation of the BRAF gene. Precision medicine is still in its early

stages, and so far it has not revealed all its potential, but it continues to grow year after year. If it keeps its promises, precision medicine could revolutionize medical health care.

Mattia Andreoletti

See also Bioinformatics; Biopsychosocial Model; Biostatistics; Evidence; Health Care Science; Infectious Disease Studies; Knowledge; Life Sciences, Contemporary

Further Readings

- Andreoletti, M., & Maugeri, P. (2019). Does medicine need philosophy? *Oral Diseases*, 25(6), 1419–1422. <https://doi.org/10.1111/odi.13143>
- Andreoletti, M., & Teira, D. (2016). Statistical evidence and the reliability of medical research. In M. Solomon, J. R. Simon, & H. Kincaid (Eds.), *The Routledge companion to philosophy of medicine* (Chapter 20). Routledge.
- Boorse, C. (1977). Health as a theoretical concept. *Philosophy of Science*, 44(4), 542–573.
- Broadbent, A. (2009). Causation and models of disease in epidemiology. *Studies in History and Philosophy of Science Part C: Studies in History and Philosophy of Biological and Biomedical Sciences*, 40(4), 302–311.
- Clarke, B., Gillies, D., Illari, P., Russo, F., & Williamson, J. (2014). Mechanisms and the evidence hierarchy. *Topoi*, 33(2), 339–360. <https://doi.org/10.1007/s11245-013-9220-9>
- Engelhardt, H. T., Jr. (1976). Ideology and etiology. *Journal of Medicine and Philosophy*, 1(3), 256–268.
- Hodson, R. (2016). Precision medicine. *Nature*, 537(7619), S49. <https://doi.org/10.1038/537s49a>
- Horwitz, A. V., & Wakefield, J. C. (2007). *The loss of sadness: How psychiatry transformed normal sorrow into depressive disorder*. Oxford University Press.
- Kingma, E. (2007). What is it to be healthy? *Analysis*, 67(2), 128–133. <https://doi.org/10.1111/j.1467-8284.2007.00662.x>
- Mackie, J. L., & Mackie, J. L. (1974). *The cement of the universe: A study of causation*. Clarendon Press.
- Parascandola, M., & Weed, D. L. (2001). Causation in epidemiology. *Journal of Epidemiology & Community Health*, 55(12), 905–912. <https://dx.doi.org/10.1136/2Fjech.55.12.905>
- Rothman, K. J., & Greenland, S. (2005). Causation and causal inference in epidemiology. *American Journal of Public Health*, 95(S1), S144–S150. <https://doi.org/10.2105/AJPH.2004.059204>

- Schaffner, K. F. (1993). *Discovery and explanation in biology and medicine*. University of Chicago Press.
- Stegenga, J. (2018). *Care and cure: An introduction to philosophy of medicine*. University of Chicago Press.
- Thagard, P. (2005). What is a medical theory? In *Studies in multidisciplinary* (Vol. 3, pp. 47–62). [https://dx.doi.org/10.1016%2FS1571-0831\(06\)80008-0](https://dx.doi.org/10.1016%2FS1571-0831(06)80008-0)
- Thompson, R. P. (2011). Causality, theories and medicine. In P. McKay Illari, F. Russo, & J. Williamson (Eds.), *Causality in the sciences* (pp. 25–44). Oxford University Press.
- Upshur, R. E. (2002). If not evidence, then what? Or does medicine really need a base? *Journal of Evaluation in Clinical Practice*, 8(2), 113–119.
- World Health Organization. (1948, April 7). *Constitution of the World Health Organization*. WHO.

MEDICINE, ENLIGHTENMENT

The 17th and 18th centuries saw important developments in natural philosophy specific to medicine, and in the work of physicians and surgeons, which formed the foundation for the scientific approach to medicine that emerged in the 19th century. Major developments in adjacent fields of research and experimentation, chiefly physics and chemistry, also proved to be of immense importance to the burgeoning fields of medicine and biology. The Enlightenment of the 17th and 18th centuries in Europe witnessed the emergence of fundamentally new ways to see the human body in terms of both physiology and pathology. Consequently, theories about what a disease is (*nosology*) and how pathological states developed (*etiology*) also changed.

However, some forms of medieval medical knowledge continued far into the Enlightenment period: humoral theory and the anatomic contributions of the Greek physician, surgeon, and philosopher Galen (129–c. 216 C.E.). Long-standing and entrenched ideas from Galenic anatomy were challenged and began to recede in the 16th century. The work of anatomists such as the Dutchman Andreas Vesalius (1514–1564) and Scotsman William Hunter (1718–1783) did much to supplant Galen's anatomical teachings, while the early nosology of English physician Thomas Sydenham (1624–1689) and subsequent disease classifications

by physiologists such as the Frenchman François Boissier de Sauvages (1706–1767) completely transformed ideas about disease and illness, away from humoral perspectives toward vitalist and mechanistic ideas of the body by the end of the 18th century. At the same time, the study of anatomy made considerable advances through the efforts of surgeons, whose practitioners grew to lay particular claim to anatomical knowledge because of its practical application in surgery.

All of these changes, spurred by the activities and writings of dispersed, and sometimes completely disconnected, experimental practices and researchers, coupled with revolutionary epistemological shifts to the questions *what is knowledge?* and *what is life?*, coalesced by the end of the Enlightenment to emerge as the scientific medicine that we are familiar with today. This entry focuses on two main elements of Enlightenment medicine: theories about the body and its health, which developed mainly with attention to the theory of humors and advances in anatomy; and classification of the body, informed by progress in surgery on the one hand and by theories and treatment of disease on the other.

Theories of Health and Body

Humoral Theory

While more inductive and scientific modes of understanding the human body started to appear in the 17th century, Galenic humoral theory persisted and maintained relevance until the later part of the 18th century. Humoral theory, which can be traced back to Greek antiquity, remained an important part of physiological thought throughout the Enlightenment period.

The idea of the four humors originated in ancient Greek pre-Socratic philosophy and was part of an overarching cosmology that sought to explain or account for all existence. The Greeks conceived of everything in the world being comprised of four elements: water, earth, wind, and fire. The writings attributed to Hippocrates extend this reductive framework to the human body and health in the form of the four humors: blood, phlegm, and yellow and black bile. These four bodily fluids were believed to underlie each individual's particular embodied experience, in health and illness, and even in our dispositions and attitudes.

Health in humoral theory centered on the idea of balance; balance in a person's diet, sleep, and exercise was reflected in an overall balance and equilibrium of the four bodily humors. Humoral balance constituted good health. If this equilibrium was broken due to either deficiency or excess of one or more of the humors, disease would ensue. For example, a tumor was understood to be an accumulation of black bile in a particular area of the body. An excess of yellow bile could cause erysipelas (red rash); a buildup of phlegm caused edema (fluid buildup).

Humoral theory continued during the Roman Empire and is most associated with the writings of the Greek physician Galen. Galen was born in Pergamon, now Bergama in Turkey, where he trained as a physician before eventually moving to Rome to work as a doctor. As a prolific researcher and writer, he rose to a high level of fame and attended to many significant figures of the age, including Emperor Marcus Aurelius (121–180 C.E.). Galen was a prolific writer; many of his books explored the topics of medicine and anatomy; others discussed philosophy and politics. A devastating fire in Rome in 191 destroyed a substantial portion of Galen's writings. Subsequently, the gradual collapse of the Western Roman Empire through the 4th and 5th centuries resulted in the disappearance from the European landscape of a large part of Galen's writings. During the medieval period, or what has traditionally been called the *Dark Ages* in European history, Hippocratic and Galenic writings survived in the Eastern (later the Byzantine) Empire and in the Arab and Persian Middle East. Numerous Arab and Persian philosophers, such as Rhazes (Abū Bakr al-Rāzī; c. 864–c. 935), Avicenna (Ibn Sina; 980–1037), and Averroes (Ibn Rušd; 1126–1198) among others, translated these writings from Greek into Arabic and spread Hippocratic, Galenic, and Aristotelian knowledge more generally. Gradually, this knowledge traveled back to Europe through Moorish Spain where it was incorporated into Christian philosophy, which became highly influenced by Greek natural philosophy, especially Aristotelian, as Christian philosophers like Thomas Aquinas (1225–1274) and William of Ockham (c. 1285–1347), incorporated it into a strict conceptual framework that would become known as *scholastic philosophy*, for its emphasis in Europe's burgeoning university culture.

Within the medieval period, the doctrine of the four humors was sometimes referred to as the four temperaments as it was believed to explain each person's particular disposition. Therefore, one could be deemed a melancholic (depressed) due to an excess of black bile or a phlegmatic (cold and tired) due to an abundance of phlegm, both humors characterized by coldness, dampness, and an overall analogous relationship to the winter season. Black bile and its association with coldness and dampness had a particular influence in concepts of pathology. These are just a few examples of how the humoral doctrine formed a complete cosmology/explanatory framework within medieval Europe for understanding the relationship between the body, health, disease, and the natural environment.

Humoral theory began to recede starting in the 17th century with the rise of experimental philosophy and the development of competing physiological frameworks, such as English anatomist William Harvey's (1578–1657) ideas about the circulation of blood (discussed below). New approaches to classifying and understanding diseases, as found in the writings of the "English Hippocrates" Thomas Sydenham, enjoyed widespread influence in subsequent developments in European medicine and pathology. As a formal medical theory, Galenism was largely discredited by 1700. Compared to the emerging approaches to natural philosophy exemplified by experimentation and systematic reasoning, Galenism was inaccurate and unreliable. It could not account for many new observations that were arising from new approaches and techniques. However, in the clinical domain, as a theory applied to the practice of diagnosis, the notion of humoral equilibrium, the *body in balance*, continued to pervade medical practices about health and illness. While famous Enlightenment doctors like Sydenham, William Cullen (1710–1790), and Giovanni Battista Morgagni (1682–1771) may have considered Galenic ideas outdated in theory, working physicians still drew from humoral theory when diagnosing and treating individual patients. Indeed, these holistic ideas of balance and equilibrium remained important physiological concepts even in 19th-century cell theory, and in Claude Bernard's *scientific physiology*, which depended on the functional interaction and balance between

numerous organs and the chemicals they produced within the body. These originally humoral ideas about balance and equilibrium persist, even if only analogically, in modern understandings of physiology such as homeostasis and cell autophagy. While the Enlightenment brought about many changes in medicine, certain core ideas survived in a different guise.

Anatomy

The traditional view within the history of science is that the prominence and authority of humoral theory and Galenic teaching began to decline during the 16th century. This shift in medical thought corresponded with widespread social upheaval throughout Europe. Notably, as Martin Luther's 95 Theses were translated into various vernacular European languages in the 1520s, Swiss-born physician Theophrastus von Hohenheim—"Paracelsus" (c. 1493–c. 1541)—staged public protests across Germany during which he reportedly burnt Latin copies of Galen and Avicenna in a dramatic show of opposition to the old order of medicine. Medicine, religion, and politics were very closely tied.

Various other significant events in natural philosophy, including the revolutions wrought by Polish astronomer Nicolaus Copernicus (1473–1543); the beginning of transatlantic exploration and colonialism following the Portuguese voyages of discovery; and the late-Renaissance humanistic ideas exemplified by the works of Leonardo Da Vinci (1451–1519) and anatomist Andreas Vesalius, all either directly challenged Galenic doctrine or the broader Christian scholastic systems of which it was part.

Centuries earlier, the discovery and assimilation of Galen's writings were themselves part of a late-medieval intellectual revolution. While certain cloistered intellectual communities would have known of Galen, Galenic doctrine was for practical purposes lost to European medicine for a millennium following the collapse of the Roman Empire in the 5th century. Galen's writings were effectively unknown in the libraries of most medieval monasteries until the cusp of the early European Renaissance. Beginning in the early 14th century, Galenic anatomical texts like *On the Use of the Parts* and *Anatomical Procedures* underwent a revival among

anatomists in European universities. First in Bologna with Mondino de'Luzzi (1275–1326) and the publication of his *Anatomica* in 1316, and then into Vesalius' time with the activities of anatomists like Jacobus Sylvius (1478–1555) and Johannes Gutiner (1510–1574) who recreated the dissections and procedures described by Galen. The point of this anatomical research was to see what Galen saw, to find what he found. Galen's authority was as yet unchallenged.

It was Vesalius who first openly questioned the accuracy of Galen's texts and even suggested that Galen himself had never dissected a human body. However, the influential Dutch anatomist represented less of a discontinuity with an ancient anatomical tradition and more of a distinct school of anatomy within an extensive field still very much dominated by Galen supporters, and those who sought to adhere to Christian doctrine. For these adherents, many of whom were at the same time trying to defend the Roman Church against Protestantism, a belief in the purity and accuracy of Galenic texts equated to the purity and accuracy of Church scripture in general. In this way, medicine and medical knowledge was very much a social discourse entrenched in religion and politics.

What was revolutionary about Vesalius was that he openly questioned the primacy of these classical texts and argued for the importance of experience, applied dissection, and observation in anatomical investigations. Galenic books ceased to represent the whole of the truth of anatomy. Even those who criticized Vesalius, like anatomist Jacobus Sylvius (1478–1555) and physician William Harvey, conceded that there were, indeed, differences between Galen's anatomical descriptions and the bodies that they themselves dissected. One creative solution to this inconsistency was to speculate that the shapes of human bodies had changed since Galen's time—that God had created humans differently in different eras. This managed to preserve the sanctity and influence of Galen's writings a bit longer, and at the same time acknowledged new anatomical observations and corrections, which could no longer be denied.

Vesalius himself found many mistakes in Galen's descriptions, and his detailed drawings, presented in his 1543 work *De Humani Corporis Fabrica* (*The Structure of the Human Body*), were

a stark departure from previous pictorial depictions of dissections that focused on representing the entire ritual of dissection, with all its participants, like the Reader who dictated Galen's instructions and observations, the dissectors dissecting the body, and the audience witnessing the spectacle.

Exhibiting much artistic skill, Vesalius's inductive naked-eye observation of the dissected body, supported by detailed true-to-nature drawings of the body and its organs, represented a stark departure from Galenic conceptions and serves as an early example of the approaches to thought and knowledge—experimentation and empiricism—that were the hallmark of Enlightenment thought and most explicitly elaborated by Francis Bacon in terms of a scientific method in 1620.

But by the 17th century, Vesalius was already outdated. While he continued to inspire as an early champion of reason and knowledge who challenged established medical authority, further additions and corrections made to anatomy after his time quickly left his actual findings and work in the archives of history. Following his death in 1564, several further developments occurred in the sphere of anatomy. For example, lacteal lymphatic vessels were discovered in the 1620s by Italian physician Gaspare Aselli (1581–1625) and the lymphatic system was more generally described in the human body by two chemists, the Swede Olaus Rudbeck (1630–1702) and the Dane Thomas Bartholin in 1650–1652.

Perhaps the most monumental anatomical achievement, symbolic of the Enlightenment era to come, was the new theory of the circulation of blood proposed by William Harvey in his *An Anatomical Disquisition on the Motion of the Heart and Blood in Animals* (1628). Whereas the precise function of the heart had been much speculated on, including the candidate theory that it acted as furnace to heat the blood, Harvey proposed that the heart, with its two ventricles, continuously pumped blood through the body using arteries and the same blood was carried to the different parts of the body. Harvey showed that it was simply impossible for the liver to continuously create the sheer volume of blood that constantly pumped and pulsated throughout the body of any given animal. As with other experimentalists or naturalists and their theories, Harvey's ideas were not

uniformly accepted across England and Europe right away. His perspective had both its supporters and critics. Predictably, the criticism focused on Harvey's deviation from Galenism. These sorts of criticisms and objections became fewer throughout the 17th century, as medicine, and natural philosophy more generally, more emphatically adopted methods of observation and induction.

The following century brought further complexity to anatomy in the form of comparative anatomy that compared the composition of different animals, including humans, in the placement and shape of organs, skeletal structure, and other features. For example, Nicholas Tulp (1593–1674), a surgeon and the mayor of Amsterdam, made some early observations on the anatomy of chimpanzees and orangutans brought from the Congo. Some of Harvey's test subjects for his theories on the circulation of the blood were also animals like eels and snails. The study of comparative anatomy and animal experimentation became an important enterprise for anatomists and natural philosophers in the 18th century. Many surgeons and physicians were skilled anatomists, but the anatomical sphere also extended into natural philosophy and early studies into animal behavior, or zoology. Georges Cuvier (1769–1832) is a worthy example of this intersection. He was both a skilled anatomist and zoologist whose work was cited and referred to by both naturalists and medical practitioners when the study of animal bodies, as an analogue to the human body, became an important avenue for research and experiment.

Classifying the Body

Surgery

Medical surgery has a peculiar history. In the present era, surgery and medicine are seen and understood to be deeply interconnected. But in the 18th century, medicine and surgery were still viewed as separate pursuits. Surgical societies and surgical colleges developed in relative isolation from long-standing anatomy- and theory-based medical schools. Beginning with medieval barber surgeons and other folk practitioners who would travel from town to town, surgery very gradually developed into a formal part of medicine.

While *proper* medicine in medieval times existed as an esoteric and philosophical pursuit for university lecturers, these medieval barber surgeons were often those to whom the common European citizen went for help. They pulled teeth, cut out ingrown toenails, helped with injuries like lacerations and broken limbs, and, owing to the ravages of syphilis, performed basic forms of rhinoplasty with skin grafts or even metal prosthetic noses.

Due to the demands of surgery and the skills required to wield a knife, many surgeons also became skilled anatomists. This was before even a rudimentary understanding of antiseptics and anesthesia, so patients often died during surgery from sepsis and shock. Postmortem examinations were commonly performed by surgeons. Anatomy as a discipline slowly drifted from the domain of formal medical doctors, closer to the domain of surgeons, which had a practical—living—component that work with cadavers in a lecture theater did not match. The anatomical knowledge that developed through the work of surgeons, combined with an increase of the practical applications of surgery to more illnesses and disorders, including lithotomies (stone removal), cataract surgeries, and herniotomies, raised the prominence of surgeons across Europe. Physicians had to understand anatomy as well, but they did not, strictly speaking, need to know those things that they did not practice. On the other hand, it became crucial for the surgeon to be a skilled anatomist at the same time as a competent physician.

Gradually, these barber surgeons began to organize themselves in guilds and professional organizations. A division began to take place between those who remained barbers practicing surgery on the side, and those who practiced surgery in a learned and focused way. The United Company of Barber Surgeons organized in London as early as 1540. Later, in 1745, this organization reformed to be the Company of Surgeons, further distancing themselves from their barber history. But the general division between English surgeons and physicians remained an important one throughout the century. However, some of the most prominent and well-known English medical authorities of the 18th century were surgeons like William Cheselden (1688–1752), Percivall Pott (1714–1788), and John Hunter (1728–1793). Cheselden worked on new methods of lithotomy

and cataract surgery. Pott not only made innovations with orthopedics and fractures but also famously established the link between scrotal cancer and the work of chimney sweeps; the first recognition of an environmental factor in the development of cancer. John Hunter, often called the *father of scientific surgery*, had a profound effect on surgical teaching in England and highlighted the importance of systematic study and surgical experimentation on animals and humans. Hunter was also a good example of a surgeon-anatomist, and his large anatomical collection of animal and human organs is housed today in the Hunterian Museum at the Royal College of Surgeons in London, England.

In France, in 1731, King Louis XIV opened the Académie Royale de Chirurgie. This institution was to mirror the surgical society in England. Two royal surgeons to the king, Charles Georges Mareschal (1658–1736) and Germain Pichault de la Martinière (1697–1783), were among its founding members. Part of this society's eventual importance was in its periodical publication, *Memoires de L'Academie Royale de Chirurgie*, which was the first periodical publication specializing in surgery. It was also in France, in part due to the many French involved conflicts in the 18th century, that surgical practices flourished as many military surgeons advanced methods of ligature of blood vessels and amputation, tendon surgery, and orthopedic surgery such as the resection of bones and joints. Famous figures of the French surgical tradition included Jean-Louis Petit (1674–1750), Henri-Francois le Dran (1685–1770), and Francois Quesnay (1694–1774), who all worked in Parisian hospitals. Like their English counterparts, these French surgeons primarily focused on the extremities and the visible illness—those maladies that were on the surface of the body or in the extremities; those parts which could be fixed in a timely manner, and without causing the patient severe pain or killing them. The abdominal cavity and the thorax still remained the purview of anatomy and dissection. It would not be until the 19th century that abdominal and thoracic surgery became possible and relatively safe for patients, and a greater understanding of the specific role of each organ, as well as their mutual interdependence, was better understood. Surgeons of the 18th century

recognized the practical limitations of their practice. Even John Hunter reportedly disliked performing surgeries precisely because of the pain that it caused patients.

Surgery, therefore, existed as a practice independent from academic, that is, theoretical, medicine for centuries. It was only toward the beginning of the 18th century that the importance of surgical practice was recognized by the public, political leaders, and medical doctors alike; it had become a vital part of medicine and a medical profession. Many practical procedures that worked to relieve pain and treat everyday problems such as broken bones, hemorrhoids, or bladder stones elevated the surgical profession and the demand for practicing surgeons. Anatomical knowledge, likewise, shifted toward the domain of the surgical profession.

Nosology and Disease Theory

Anatomical knowledge and its development were important for the entire medical domain. Yet most anatomical knowledge had come from working with corpses, and looking at an organ in the body, especially that of a cadaver, did not automatically translate into knowledge about the workings of that organ or its role in maintaining the life of the organism. Furthermore, it could say nothing about a disease affecting the organ. To understand what disease was, to differentiate it from symptoms, and to understand how the disease became clinically manifest in terms of its symptoms, required the development of ideas and theories about disease and pathology. These theories of health and disease began with the development of classification systems, not just in the sphere of medicine but within the entire field of natural philosophy of the 17th century.

The explosion of transnational trade and exploration in the 16th and 17th centuries resulted in many private collections of different objects, trinkets, plants, creatures, fossils, and even humans, from distant corners of the world. These collections, usually commissioned by wealthy individuals, eventually led to the practical need to organize and classify the objects in a coherent and logical manner. Some of these collections were modest and contained within a *Wunderkammer* (cabinet of curiosity) or they could be large and require an entire museum, such as the collection of Russian

Tsar Peter the Great, still housed at the Kunstkammer in St. Petersburg. These various collections became the object of study for many natural philosophers. The most celebrated of these early classification systems was that of the Swedish botanist, zoologist, and physician Carl Linnaeus (1707–1778), who published several major works, including *Systema Naturae* (1735). There were other classification frameworks besides that of Linnaeus, such as the geographic analysis of species by Eberhard A. W. Zimmermann (1743–1815) and George-Louis Leclerc Buffon's (1707–1788) classification studies first published in 1749. These competing ideas of classification and taxonomy were a precursor to the evolutionary theories of the 19th century. In the 17th century, the Italian *Accademia dei Lincei* (Lincean Academy) was the first to attempt to create an all-encompassing theory of order and existence for the simple honeybee in honor of their former member Cardinal Maffeo Barberini, who was then Pope Urban VIII. Interestingly, Galileo Galilei (1564–1642) was another prominent member of the Academy. These observations and classifications of the honeybee were first published in 1625 as the *Apiarium* and made use of the recently invented microscope to observe the bees in detail never before seen.

The same sort of attempts at taxonomical classification also started to appear in the medical and anatomical sphere. Perhaps the most famous 17th-century example was Sydenham's classifications of fevers. He published his observations and a classification of fevers and febrile disease in 1666 as *The Method of Curing Fevers*. His classification focused not only on the different types of fever but also grouped together different types of symptoms according to their compatibility and likeliness to occur together. Sydenham wrote down case histories and used large samples of patients to form his conclusions. While his observations on disease were still heavily reliant on humoral theory and were soon replaced by others, the general framework he pioneered was widely adopted to understand diseases and their symptoms. The same idea of disease as a combination of symptoms was still important at the end of the 18th century, even if by then, anatomy and the morbidity of specific internal organs was also considered when determining or classifying disease.

In 1731, Montpellier physician François Boissier de Sauvages (1706–1767) followed in the tradition of Sydenham to create a modernized version of a disease nosology that still primarily focused on symptoms to identify disease but implemented a system based in the categories of *classes*, *orders*, *genera*, and *species*—his classification system encompassing over 2,400 species of illnesses. The English physician William Cullen likewise worked on his own classification scheme and published it as *First Lines of the Practice of Physic* in 1778. Interestingly, Cullen doubted the utility of anatomical knowledge in diagnosis and noted that observations of internal organs and lesions were too unreliable and inconsistent to use in diagnosis and classification.

This medical tradition and theory viewed the body in a holistic sense and its vitality or health as a finite quantity. Before the late 19th century and investigations into the possible existence of specific pathogens that caused disease, the medical concept of *disease* was thought of in a more fluid manner. Life hung in a balance of health and illness and once the process of illness began, the person would gradually decline in health until their inevitable end. This process could take months or years, it could stall, and then proceed rapidly, as the disease shifted from one unpredictable form to another. In a medical climate where a large part of pathology was speculative, the main source of knowledge about the illness came from the patient and the symptoms they reported. Only the patient could feel the pain in their side or their chest, and so the external, that is visible or empirical, symptoms played a major part in a physician's determining and classifying the illness. Within this disease framework, a simple cough could turn into a fever, a fever into tuberculosis, and tuberculosis could take years to transform into cancer, signaling the end of life. Instead of separate diseases and illnesses, there was one essential *disease force*, which depleted life and which took on many different forms.

Life was likewise seen as an essential force in the vitalist tradition. Anatomist and early embryologist Johann Friedrich Wolff (1733–1794) called this force *vis essentialis* and ascribed to it all manner of functions. This vital force was why the embryo could grow into an adult animal, and it was also why a wound would heal, or why a broken

bone would grow back together. The vitalist ideas altogether became very prominent among German practitioners and naturalists. At the universities of Göttingen and Halle, figures like Wolff, Georg Ernst Stahl (1659–1734), and Johann Friedrich Blumenbach (1752–1840) embraced these vitalist principles and incorporated them into their studies of embryology, physiology, and even early histology and cell theory. These German doctors would form the intellectual tradition of *Naturalphilosophie*, which was an important school of thought in 19th-century German biology.

The major challenge to this vitalist viewpoint came from the school of mechanistic thought. This tradition was inspired by the philosophy of earlier luminaries such as René Descartes (1596–1650) and Isaac Newton (1643–1727). Within this theory, the body was seen as a machine, like a clock with gears and springs, that functioned according to physical and mathematical laws. Descartes made some observations about the mechanistic nature of the body in terms of the function of muscles and the circulation of blood in his unfinished work *L'Homme* as early as the 1630s. The first medical practitioner to be a proponent of this mechanistic theory of the human body was Hermann Boerhaave (1668–1738), a Dutch physician who worked in particular on problems of inflammation and performed some of the first analyses of tissue using a microscope. Within his mechanistic framework, the body functioned according to the laws of hydrostatic forces and pressure. Disease was described as the disequilibrium of these pressures within the body. Inflammation caused by blunt force, for example, was a build-up of excess pressure in the affected part of the body. Boerhaave's ideas were published in 1708's *Institutions of Medicine*, which became a highly influential book across Europe. Later vitalists would challenge some of Boerhaave's ideas. This included one of his own students, Albrecht von Haller (1708–1777). Haller, who was an intellectual opponent to the embryological theories of Caspar F. Wolff (1733–1794), also modified the hydrostatic vitalism of Boerhaave, instead focusing on the importance of the nervous system in the body's functioning.

Other researchers continued to pursue some of these hydrostatic mechanistic ideas. One outcome was the creation of the *blastema-exudation*

hypothesis toward the end of the century. In this hypothesis, the fluids of the body pooled in the injured or damaged areas and then congealed and solidified. This hypothesis was used by a variety of authors to explain the appearance of different pathological conditions; everything from inflammation, *a la* Boerhaave, to the appearance of ulcers, cysts, and even tumors. This perspective remained active into the next century. The downfall of many mechanistic and vitalist ideas did not occur until the mid-19th century development of cell theory. Cell division in plants was first observed by Matthias Schleiden (1804–1881) and Theodor Schwann (1810–1882) in the 1830s. These principles were then applied to the human body by pathologists, notably Johannes Peter Müller (1801–1858) and Rudolf Virchow (1821–1902). Mechanistic and vitalist perspectives were abandoned in favor of the analysis of the cell, echoing Virchow's dictum, *Omnis cellula e cellula* (All cells come from similar cells). The bacteriological revolution at the end of the 19th century brought even more stark changes to the medical and biological sphere of knowledge. As a result, medicine at the end of the 19th century looked nothing like medicine at the end of the 18th century.

Conclusion

As the preceding sections have shown, a single, unitary concept of *medicine* such as we have today did not exist before the 19th century. This concept of unified medicine came about because an assemblage of medical practitioners, anatomists, surgeons, and natural philosophers gradually built up a body of knowledge about the living body and its relation to the natural world. This process was not coordinated, and important developments could come from the unlikely of places.

As we have seen with the rejection of Galenism and humoral theory, this process was also a cultural and political matter. And was the case with the continuation of humoralism into the 18th century, the process of rejecting *old* knowledge wasn't straightforward or even always successful. Perhaps what best describes the medical achievements of the Enlightenment is the creation of new theories. Enlightenment-era medical practitioners, and other thinkers, were not able to learn to cure and effectively treat many of the same diseases and illnesses that had plagued Europe for

centuries, but they did begin to develop new ways of seeing the living body, and new ways of seeing disease and illness, that fundamentally changed the concepts of life and existence.

Dimitry Zakharov

See also Biology, Evolutionary; Cell Theory; Empiricism; Laws of Nature; Medicine, 19th Century; Medicine, Antiquity; Medicine, Middle Ages; Medicine, Renaissance; Scientific Revolutions

Further Readings

- Bognon-Kuss, C., & Wolfe, C. T. (Eds.). (2019). *Philosophy of biology before biology*. Routledge.
- Bynum, W. F. (1994). *Science and the practice of medicine in the nineteenth century*. Cambridge University Press.
- Canguilhem, G. (2010/1965). *Knowledge of life*. Fordham University Press.
- Coleman, W. (1964). *Georges Cuvier, zoologist*. Harvard University Press.
- Cunningham, A. (1997). *The anatomical renaissance: The resurrection of the anatomical projects of the ancients*. Ashgate.
- Daston, L., & Galison, G. (2007). *Objectivity*. Zone Books.
- De Moulin, D. (1983). *A short history of breast cancer*. Kluwer Academic Publishers.
- Duffin, J. (2010). *History of medicine: A scandalously short introduction*. University of Toronto Press.
- Foucault, M. (2010/1963). *Birth of the clinic*. Routledge.
- Freedberg, D. (2002). *The eye of the lynx*. University of Chicago Press.
- Golinski, J. (1998). *Making natural knowledge: Constructivism and the history of science*. University of Chicago Press.
- Hagner, M. (1999). Enlightened monsters. In W. Clark, J. Golinski, & S. Schaffer (Eds.), *The sciences in enlightened Europe* (pp. 175–217). University of Chicago Press.
- Harrison, M. (2004). *Disease and the modern world*. Polity Press.
- Jewson, N. D. (1976). The disappearance of the sick-man from medical cosmology, 1770-1870. *Sociology*, 10(2), 225–244.
- Porter, R. (1997). *The greatest benefit to mankind*. Fontana Press.
- Temkin, O. (1951). The role of surgery in the rise of modern medical thought. *Bulletin of the History of Medicine*, 25(3), 248–259.

MEDICINE, MIDDLE AGES

Generally during the Middle Ages, extending from about 500 to 1400–1500 C.E., people in Europe still followed the concepts taught by the ancient Roman physician Galen (129–ca. 210 C.E.), whereby a body was made up of four *humors*, or fluid substances, and illnesses were a result of these not being in proper balance. Some Christian theologians in Europe argued, however, that disease was often a punishment for sin or for divine disfavor. They urged that people who were suffering should pray and endure these afflictions in much the same manner as Job had done in the Hebrew Bible. Cure came through the intervention of God, and therefore it was necessary to continue to pray. Although Gregory of Tours during the 6th century C.E. includes some of these views in his writing, they certainly predate him.

This entry explores the development of medicine during the Middle Ages in Europe, in the Arab world, and in China and the Khmer Empire (present-day Cambodia). The influence of ancient Greco-Roman sources such as Hippocrates and Galen is discussed, as is the work of early Christian physicians and the widely noted, pioneering Arab physician Avicenna, whose works were translated into Latin. The response to lethal infectious disease, chiefly the plague, or Black Death that devastated cities and entire regions in the 14th century, is described.

Beginnings and Early Practices in Europe

Some of the early Christian writers point to miracles that had been documented. Jesus of Nazareth had been involved in curing the lame, and treating the dumb, the blind, a leper, and a man who had already died. His disciple Paul of Tarsus was able to treat the father of Publius (later St. Publius) in Malta after having been shipwrecked on the island. The father had been suffering from dysentery and was cured.

A medieval manuscript in the British Library includes references to the Anglo-Saxons in England conducting some fairly crude surgery. Much of their beliefs in illness was that bleeding a wound was a cure for many ailments. This is mentioned

when a physician called Cynefrid had the task of operating on Queen Etheldrida of East Anglia in 660. The daughter of King Ine, she had a swelling under her jaw on her neck, and the doctor, according to Bede in his *Ecclesiastical History of England*, lanced the abscess. She died following this, with Bede reporting that she believed that the illness had been caused by God because of her vanity over her neck. It was reported that when her body was exhumed some 16 years later, it was found that her body was preserved and the neck wound had gone.

There are many other stories of miraculous cures in Anglo-Saxon England. After the murder of King Edward the Martyr in 978, his body was concealed in the house of a blind woman. During the night, the woman regained her sight, and when Edward was buried at Shaftesbury Abbey, there were reports of blind people coming to pray at his grave. He was subsequently declared a saint.

Although the Greeks commonly believed in washing wounds, the Roman writer Pliny the Elder suggested that the wounds could be closed using honey, and this was a practice followed by the Anglo-Saxons, who did not follow what would now be seen as basic customs of hygiene. Indeed, there seem to have been regular problems with the fouling of water, and the Saxons also often buried people inside churches and scattered the dust from these burial places on water, which would then be given to people and cattle suffering from ailments in the hope of a cure.

In the early Middle Ages, there was a major rise in the population of England, France, and many other parts of Europe. At the same time, infant mortality was high, and many babies were extensively swaddled as a way of trying to prevent rickets. During this period, what emerged as a major issue was chronic inbreeding in many villages. The Bible had an injunction against the marriage of first cousins, and this had served Christian Europe well. But with the rise in population came many second cousins marrying, thereby producing a wide range of associated health issues.

Although most people with illness were treated in their homes, usually by family members, monasteries and abbeys were emerging as places where people could seek treatment. This has its origins in the emergence of the importance of Christianity as a religion in the Roman Empire, with Pantaleon of

Nicomedia (d. ca. 305), practicing medicine among the poor, and Alexander of Tralles (525–605 C.E.) working as a skilled physician and emphasizing the powers of observation. St. Benedict of Nursia (d. 547 C.E.) entreated his monks to care for the sick, but he also instructed them to refrain from reading about medicine because of his belief in divine intervention as the only certain cure.

Some monks had the role of being herbalists, as did as the fictional Cadfael in the novels of Ellis Peters (the pen name of Edith Pargeter, 1915–1995). There is a reference to Baldwin, a French monk from the Abbey of St Denis coming to England to treat Edward the Confessor and later being made the abbot of Bury St. Edmund's. He continued to treat William I "The Conqueror," and after his death in 1098, Faricius of Abingdon, another abbot who had been trained in medicine at Salerno in Italy—the only major school of medicine in Europe at that time—took over that position as the court physician and was responsible for the treatment of many court officials. Abbot Warin of St Albans was another monk who was renowned for his skills.

It was at Salerno and then also at Montpellier in the south of France that the works of the Greek writers Hippocrates (ca. 460–ca. 360 B.C.E.) and Galen of Pergamon were translated and their references to the different bodily humors were treated deferentially. It was long recognized, however, that there were limits to medical knowledge. Since ancient times, lepers had been cast out from communities, and this continued during the Middle Ages, with leper *colonies* throughout medieval Europe. The major change was that whereas in antiquity this was regarded as an affliction that may have come about through the actions of the sufferers, in the Middle Ages many people came to pity the lepers, with some wealthy people donating money to build churches and almshouses for the lepers—prayer being seen as the main way in which they might be cured.

Medicine in the Arab World

Most treatments in medieval Europe consisted of herbal medicines and balms, as well as prayer, visits to religious shrines, and taking part in pilgrimages. Some of the herbal cures were relatively effective at

dealing with minor ailments, although the placebo effect cannot be underestimated. With the Crusades, the physicians of Western Europe came into contact with the work of Arab physicians such as Avicenna (Ibn Sina; 980–1037), whose works were translated into Latin. There was also much interaction between Moors and Christians in Spain—Morocco having emerged as a major center of learning in the Arab world although, interestingly, many of the successful practitioners there were Jewish or even Christian. Many Muslim doctors were trained in Alexandria and Baghdad.

Arab medicine has long emphasized cleanliness, and much greater hygiene than had been the case in Western Europe. This was reinforced by Islam with people washing themselves before prayer, and wealthy Muslims often leaving money for the building of bath houses. Perhaps because of the warmer climate in many parts of the Arab world, cleanliness was seen as much more essential than it was in northern Europe, and the European world learned much from the Arab world.

Avicenna was the author of *The Book of Healing* and *The Canon of Medicine*, which between them transformed much of the thinking about medicine in medieval times, with both remaining in use until the mid-17th century. Avicenna had managed to study the works of the Greeks, the Romans, and also the Persians and Indians. He had been influenced as a young man by philosophy, especially the *Metaphysics* of Aristotle. When he was 16 he had decided to turn to medicine and rapidly had much success treating patients. He then was appointed physician to Emir Nuh II which allowed him access to the important medical texts held by the court. His main new theory concerned the cause of epidemics and contagious diseases, which he felt came from bad *vapors* in the air and that hygiene was very important, as well as isolation of patients. He also was to support the view of Aristotle that tuberculosis was contagious. This was a view that had been rejected by Hippocrates and hence rejected by many people in Europe. As for the testing of medicines, he believed that the apparent success of some of these was often owing to chance, and felt it was important that all medicines should be tested on humans, because the effect it might have on an animal might be quite different.

The Plague and Its Aftermath

The work of Avicenna had much influence on some thinkers in Europe, and moves were made to try to prevent pollution of water supplies, although these were often fairly rudimentary in medieval cities, with the ever-present proximity of animals and humans in both cities and in the countryside adding to some of the problems. Moorish and Arab doctors—as well as Jewish doctors—started practicing in cities in Western Europe with some success. They often decried the heavy surgery that was still resorted to in order to treat many ailments in medieval Europe. Leg ailments were often treated by bleeding the wound and then sometimes by amputation. The lack of hygiene of some of the medical equipment presented problems of its own, and crowd scenes pictured in medieval manuscripts regularly include at least one in 10 people depicted as being crippled.

The plague had long been viewed with great fear since Roman times, and in 1346–1353, the Black Death struck Europe. Modern scholars believe that the plague arrived from Central Asia along the Silk Road and reached the Crimea in around 1343. It then traveled westward, initially brought by Genoese ships that traded in the Black Sea, and over the next 10 years it devastated communities and resulted in the deaths of approximately one third of the population of Europe and parts of the Middle East, although some modern scholars have thought that this was an underestimate, and some cities such as London, Hamburg and Bremen lost close to two thirds of their population.

The rapid spread of the plague and the vast number of people who succumbed led to many widely varying theories about its cause. The speed of transmittal of the disease meant that there was hardly any time for rational reflection, and with so many doctors and others dying, the Black Death severely tested existing medical theories. Although many recognized that this was a medical issue, it was not long before scapegoats were identified. Lepers were accused of spreading the disease, and there were also many claims that Jews had poisoned the water supplies, resulting in massacres of many Jewish communities such as those in Cologne and Mainz.

What was clear to many people, even from the start of the Black Death, was that it was caused by

contagion. The outbreak in Sicily could be dated to have occurred straight after the arrival of some Genoese galleys—Genoa having colonies in the Black Sea. The outbreak in Norway can be dated to have taken place following the arrival of a single ship at Askøy in 1349. Some cities refused to accept ships from Genoa or from affected areas, and a few isolated areas managed to escape the plague, notably the kingdom of Poland, the Basque areas of northern Spain and southern France, and curiously, some farming areas in modern-day Netherlands and Belgium, as well as the city of Milan. Some of these areas—such as Poland—escaped the plague because they were not on the major trade routes, especially by sea. In Milan, the city officials embarked on a policy of isolating any sufferers. If anybody showed any signs of infection, their house and the three next to it on either side would be closed up and walled up with the occupants inside. It was a drastic measure, but it saved the city from the massive number of deaths that destroyed so many other communities. In Prague, there was a belief that prayer in the city and a statue of the infant Jesus had protected the population.

There were major social changes following the Black Death. The loss of so many people in a large number of European and Middle Eastern cities reduced the amount of pollution and effluence, with consequent changes in people's lifestyles. It also had the result of massively reducing the tax base of many countries, leading to social upheavals such as the Peasants' Revolt in England. Many of the people who survived the plague had clearly been exposed to it and had built up a resistance to it. Therefore, while there were subsequent outbreaks of the plague over succeeding years, these were never as devastating as the Black Death. In the waning years of the Middle Ages, while changes in European thought concerning disease and illness led to continued advances in medicine, prayer still remained one of the major sources of solace for many people who were ill, or who had friends or family members suffering from ailments.

China and Southeast Asia

In China, fundamental concepts of medicine remained much the same, with the forces of Yin and Yang, seen as representing the dual nature of

the body, had to be brought into harmony for a cure to be effected. The use of herbal remedies and acupuncture continued to yield positive results, and there was a far greater body of knowledge recorded about these during the Tang and Song dynasties, thus helping doctors to develop more accurate ways of treating illness. Chinese doctors started to practice outside China itself. And Zhou Daguan (Chou Ta-kuan), a Yuan dynasty envoy in 1296–1297 to the Khmer Empire city of Angkor (in what is now Cambodia), recorded his surprise to see that lepers were not shunned by society there, as in China but lived in the community. He also noted the widely practiced custom of submerging oneself in water as a cure. This occurred following the reign of the Khmer king Jayavarman VII (1125–1218) who had transformed the state from a Hindu one to a Buddhist one, building large numbers of hospitals and hospital complexes throughout his kingdom.

Justin Corfield

See also Medicine, 19th Century; Medicine, Antiquity; Medicine, Contemporary; Medicine, Enlightenment; Medicine, Renaissance

Further Readings

- Billr, P., & Ziegler, J. (2001). *Religion and medicine in the Middle Ages*. York Medieval Press.
- Campbell, D. (1926). *Arabian medicine and its influence on the Middle Ages*. K. Paul, Trench, Trubner & Co.
- Dawson, I. (2005). *Medicine in the Middle Ages*. Enchanted Lion Books.
- French, R. K. (2003). *Medicine before science: The rational and learned doctor from the Middle Ages to the Enlightenment*. Cambridge University Press.
- Getz, F. M. (1998). *Medicine in the English Middle Ages*. Princeton University Press.
- Gordon, B. L., & Banov, L. (1959). *Medieval and Renaissance medicine*. Philosophical Library.
- Horrox, R. (1994). *The Black Death*. Manchester University Press and St. Martin's Press.
- Jacquart, D., & Thomasset, C. A. (1988). *Sexuality and medicine in the Middle Ages*. Princeton University Press.
- Lindberg, D. C. (1978). *Science in the Middle Ages*. University of Chicago Press.
- Macdonald, F. (2006). *The plague and medicine in the Middle Ages*. World Almanac Library.

Siraisi, N. G. (1990). *Medieval and early Renaissance medicine: An introduction to knowledge and practice*. University of Chicago Press.

Zahler, D. (2009). *The Black Death*. Twenty-First Century Books.

Ziegler, P. (1969). *The Black Death*. John Day Co.

MEDICINE, 19TH CENTURY

During the 19th century, Western medicine underwent significant changes. It did so by incorporating new scientific theories and concepts of disease. The century can be fairly evenly divided into thirds, denoting the end of Enlightenment medicine, or the age of systems, by the 1820s; the clinical medicine of mid-century; and the laboratory-based medicine of the late 19th century. These dates shift somewhat based on geography but are a good general shorthand. Historians generally have viewed the century in terms of medicine's transition from *traditional* to *modern*. This is a useful framework for thinking about the broad changes across the profession. In this context, *modernity* included the increasing importance of institutions like the hospital and asylum, new therapeutics, and (in many locations) greater authority of medical professionals (especially physicians). Medical theory specifically in the 19th century underwent significant change and helped shape many modern concepts from germ theory to genetics. Over the course of the century, regular or allopathic medicine became more self-consciously scientific, setting the stage for 20th-century biomedicine and medical specialization. *Alternative* medicine also grew in popularity, often appealing to older more system-based or holistic concepts of the human body. This entry is organized chronologically, according to into the three eras just described.

The Age of Systems

Working physicians at the turn of the 19th century faced a medical landscape dominated by competing ideas about the physiology of human bodies and how physiology could inform therapeutics. Despite that variation, most assumed that the nervous system, circulatory system, and other

combinations of organs regulated the functions of the human body as a whole system. Rather than view illness as an external invader, most doctors defined ailments as malfunctions of the body. When normal physiological functions struggled or failed as a result of internal or external strain, disease was the result. These physicians (like their predecessors for centuries before them) inhabited a world where disease was an ever-present feature of their lives, both the regular collection of chronic ailments, injuries, and childhood ills and, increasingly, threats from epidemic disorders. Patients demanded relief and understanding, which stemmed from the same mix of classical and Enlightenment roots as their physicians' educations. Family members, patients, and physicians alike looked to alterations in the body's immediate environment for the roots of distress and disorder. The systemic changes were considered the heart of the matter with respect to disease. Few ailments—smallpox, plague, and yellow fever were the exceptions to the rule—earned the designation of *contagious* in any modern sense of the word, and the search for specific disease-causing agents or toxins was limited. Similarly, therapeutics were general and drugs were understood in terms of their actions on the body. As late as the mid-1830s, one Ohio physician could argue confidently that medical philosophy *abhorred* a specific therapy—a single drug useful for a specific disease. Rather than a search for specific treatments for specific ailments, these physicians sought the best system with which to understand the human body.

The start of the 19th century coincided with the decline of the age of medical systems whose origins lay in antiquity and whose zenith occurred in the Enlightenment era (18th century). For over a century, Western medicine had relied on succession of medical and physiological systems to explain sickness, health, and the function of human bodies, incorporating new science into old frameworks. European medical theorists such as Georg Stahl, Hermann Boerhaave, and William Cullen each built on and critiqued the work of others to better articulate a complete understanding of what ailments existed and how they affected bodies. Some reached back to the humoral system of antiquity, while others picked up new scientific ideas from chemistry, hydrostatics, and physics to

try to explain the phenomena of life. Electricity, for example, proved a popular therapeutic agent as well as explanation for the phenomena of nerve communication and muscular contraction. The physical reality of life itself was thought by many to be a force or fluid influenced by a body's environment. Luigi Galvani's late 18th-century experiments with frog muscles—published in *De viribus electricitatis in motu musculari* (1792) and other texts—proved especially influential. Instrument-makers marketed electrical machines with explicitly medical uses during this period. Such ideas even captured the popular imagination as in Mary Shelley's novel *Frankenstein*, first published in 1818. The tale embodies the era's concern with the nature of life and how it could be manipulated by men of science in ways reminiscent of Galvani's work with frogs. However, whether they looked back to ancient Greece and Rome (Hippocrates, 460–370 B.C.E.; Galen, 129–c. 216 C.E.) or forward to the discovery of new elements or application of electricity in medicine, these theorists and innovators did have one important feature in common: they looked at the body as a unit.

We can see the practical implications of this view by considering the basic example of naming disorders. Tuberculosis will appear more than once in this entry and for good reason—it was especially prevalent during the 19th century—but we should start by thinking about the word itself. The term *tuberculosis* is a product of 19th-century medical ideas. It denotes a disease associated with a specific physical deformity; the tubercles or lesions that form in diseased lungs. Clinical medicine helped physicians connect these lesions to the signs and symptoms of disease they saw in living patients. Several decades later, bacteriology zoomed in more closely and connected the lesions with the presence of a specific microorganism: *Mycobacterium tuberculosis*. At the beginning of the century, however, medical doctors looked to the whole body of a single patient. They considered the patient's age, lifestyle, and physical environment to adjust their expectations and diagnosis. The disease that would come to be named tuberculosis was usually called *consumption*, a term evocative of what it actually did to human bodies. Those who suffered from the presence of *Mycobacterium tuberculosis* in their lungs and tissues appeared to physically diminish. The disease

consumed them in the sense that they appeared to disappear by losing weight and energy. In short, at the beginning of the 19th century, tuberculosis was a disease of the whole body, and by the end of the century it was understood to be an infection caused by a specific microorganism most closely associated with a specific bodily organ. Other 18th-century terms such as *fever* or *cholera* did not disappear but radically transformed in their meaning from a full-body disease and form of diarrhea to the sign of elevated temperature and a specific ailment caused by a specific bacterium, respectively.

Ultimately, in the age of systems, human bodies were permeable and exposed to the elements. Doctors and patients alike understood that life persisted through a careful balance of stimulation and rest, food and drink, and a healthful climate. Illness for most of these systems was interpreted as some sort of disruption of that careful balance that governed life processes. Ideas of exactly how that balancing act worked differed from practitioner to practitioner. The progress of medicine as it appeared from the perspective of 1800 was a story of one system proposed, proved wanting based on new observations and evidence, and subsequently replaced. That is how Scottish physician William Cullen understood the recent past in the preface to his medical textbook *First Lines on the Practice of Physic*, first published in 1782. Cullen and his mid-18th-century system continued to have relevance into the early decades of the 19th century. Some of this relevance came from critiques of his definitions of disease or focus on the nervous system by Cullen's former students. Related to systems was the persistent early 19th-century appeal to the work of Hippocrates, especially the works included in *Airs, Waters, Places*, which informed readers that their environments shaped their constitutions and therefore their health. From the perspective of systems and the influence of Hippocratic and Neo-Hippocratic medicine, it is easy to see how new spaces—especially those with unstable or unusual weather from a European perspective—would garner medical attention. If the comparatively minimal challenges of choosing the correct side of a hill to live on had medical consequences, then certainly the movement across continents and oceans did as well. This lasted well into the 19th century in the form of travel guides

and medical geographies like Daniel Drake's *Systematic Treatise on the Principal Disease of the Interior Valley of North America*, published in the early 1850s, which detailed the diseases found in the Mississippi Valley in geographic and environmental terms.

With this context, it is important to keep in mind that the changes in 19th-century medicine also took place during an era of imperialism. Imperial expansion in the global south directed research agendas and defined some diseases as distinctly *tropical*. The broad understanding of *warm* and *hot* climates with specific regional variation from the turn of the 19th century started to break down and form new definitions. The *tropics* and *tropical* illness formed as categories alongside widespread colonial encounters with Africa, Central and South America, and Asia in the mid-19th century. New technologies—including the ability to synthesize the malarial prophylactic quinine—allowed greater exploitation of regions previously considered unsafe or unsuited to European bodies. Medicine also played a role in othering and racializing colonial people. The customs and very bodies of colonized populations were pathologized and studied by doctors from the respective metropolises. As early as the 17th century, colonial practice and observation were a way for young physicians to make a name for themselves. While the tropics themselves became a new space for intellectual inquiry, the people of tropical spaces and their habits became subject to medical scrutiny as well. In a variety of colonial contexts, colonized subjects were recast as ill-suited to caring for their own health. While the presumed physical challenges of tropical climates were proposed as stimulating and helpful in promoting White masculinity, those same spaces were cast as detrimental to colonized populations.

The Role of the Clinic

While systems dominated the Anglophone medical world until the 1820s, the winds of change blew in France starting in the 1790s, where revolution affected medicine as well as politics. New ideas about the value of citizens and radical equality of bodies altered the way physicians were able to understand disease in their patients beyond Paris.

The clinical medicine that arose from the turmoil of the French Revolution provided a counterweight to the more holistic ideas of the past. Although observation and empiricism had been part of medicine in the previous era, clinical methods centered such observation in a different way, one where it drove theory and increasingly relied upon quantification. A new take on empirical observation and ideological commitments to the equality and *sameness* of bodies combined to shift medical analysis from the whole body and unique patient to the disease and diseased organs within a generic ailing body. This new concept was well suited to the major medical concerns of the era and the waves of epidemic disease that characterized mid-century society. Whether yellow fever, tuberculosis, or cholera, interchangeable bodies helped doctors and governments put the emphasis of medical analysis on disease rather than on individuals. Furthermore, a sense of therapeutic nihilism or pessimism set in when the techniques of the past did not meet the challenges of 19th-century healthcare.

The *clinic* of clinical medicine centered on the modern hospital. European capitals, including Paris, were home to large hospitals by the turn of the 19th century that could easily house thousands of patients organized by gender and disease type. While such institutions often had medieval roots, the separation of patients based on disease type and greater intervention of medical professionals during the early modern period slowly transformed them into something recognizable as a hospital today. However, the patients in early 19th-century hospitals in Europe and North America were almost exclusively the poor, and the hospitals were run by municipalities or religious organizations. Anyone who could afford to see a doctor at home arranged a house call. Those *worthy poor* who might appear at an institution like Philadelphia's Pennsylvania Hospital could expect treatment free of financial charge but were also expected to show gratitude and serve as living pedagogical tools for medical students. The implied exchange gave physicians and their pupils access to numerous patients and the tacit ability to test new techniques on such patients and provide basic healthcare to a subset of the poorest in the community.

At the same time, however, the role of the public in maintaining health gained renewed attention.

The metaphor of the *body politic* long predated the 19th century, reaching its most iconic image in Thomas Hobbes's *Leviathan*, first published in 1651. In the text the English philosopher framed his work stating, "by Art is created that great LEVIATHAN called a COMMON-WEALTH, or STATE . . . which is but an Artificial Man [*sic*] . . . for whose protection and defense it was intended" (Hobbes, Introduction, pp. 3–4). Nevertheless, modernity instilled a new urgency for considering the connections between the health of a people and the health of a state. Republics in particular felt the need to prove the health of the state through a demonstration of the health of its citizens. This encouraged a flurry of discussions and re-assessments of what historian George Rosen calls *social health*. The social contract between citizens and their government could be extended to include the health and safety of those citizens. Meanwhile, republican physicians like Benjamin Rush in the United States argued that good governance itself required a healthy and educated citizenry, thus placing medicine at the center of 19th-century politics and demanded public responses of healthcare. Caring for the sick and *worthy* poor as well as preventing epidemics took on new political meaning during the Age of Revolutions. The massive Parisian hospitals were an ideal location to put these ideas into practice and reframe care of the poor into care for the state itself. Rather than center individuals and their unique physiological systems as the unit of analysis, clinical physicians isolated the disease within the bodies of many citizens.

In the United States, pre-revolutionary institutions like the Pennsylvania Hospital and medical faculty in the same city were joined by similar institutions throughout the country as well as new efforts to distribute healthcare, such as the dispensary. Early 19th-century dispensaries acted as charity outpatient clinics where the poor could seek medical advice and treatment before they became sick enough to require a hospital. Doctors volunteered time or took low pay in dispensaries and hospitals in exchange for access to patient bodies. Institutions allowed practitioners to see a greater number of patients than could be seen by house call. Additionally, the hospital in particular allowed for the comparison and sorting of patients by disease type, age, and sex. In France, this sorting

and organization allowed doctors to see the sick body in a new way—not as an interconnected whole but as a unit damaged or infected in a way that could be replicated in others.

As noted earlier, tuberculosis is the classic disease to discuss in the new clinical context. Consumptive patients arrived in hospitals in staggering numbers and often terminally ill. The ailment was widespread among Europeans and North Americans during the first half of the 19th century regardless of social class. Other diseases soon followed suit. Physical manifestations of illness like the Peyer's patches of typhoid fever joined a growing number of lesions that located illness in specific parts of a generic human body. Physicians such as the Englishman Thomas Addison (1793–1860) started describing the *type* of patient associated with a specific disease and sought explanations in a specific organ system.

Another example of this shift in perspective can be seen using the example of new technology, in this case the stethoscope. French physician Rene-Theophile-Hyacinthe Laennec (1781–1826) began his discussion of auscultation, writing:

In 1816 I was consulted by a young woman presenting general symptoms of disease of the heart . . . Taking a sheaf of paper I rolled it into a very tight roll, one end of which I placed over the praecordial region, whilst I put my ear to the other. I was both surprised and gratified at being able to hear the beating of the heart with much greater clearness and directness than I had ever done before by direct application of my ear. (quoted in Bishop, 1980, p. 448)

This famous incident has been presented as the advent of the modern stethoscope in medicine. While the curling of a paper tube is a far cry from modern means of reading and understanding the sounds of internal organs, it does show the kind of innovation and change characteristic of the 19th century. After his success with the private patient, Laennec did what all good clinical physicians of his generation did and tried to apply the principle to patients in the hospital.

In addition to Laennec, a series of French clinicians took advantage of hospital practice and access to bodies to shape new ways of looking at

illness. Pierre-Charles-Alexandre Louis (1787–1872) shifted the clinical gaze to focus on signs of disease rather than symptoms. Whereas a symptom denotes how a patient feels—a key part of 18th-century physician–patient dialog—a sign does not require the same interaction or relationship with the patient. Signs are observable and often measurable changes in the body that point to a particular ailment. Think of the characteristic rash of chicken pox or the high temperatures of fever. Louis's focus on the lesions left on lungs by tuberculosis turning the vague ailment of *consumption* into something that could be readily observed—if only in dead bodies at first. Looking at bodies in aggregate allowed physicians to view disease in a new way. As noted by historians like Christopher Hamlin and Charles Rosenberg, terms like *fever* went from general disease terms to understand ill states of human bodies to signs of illness. Clinical work also saved lives. The observations of the physician Ignaz Philipp Semmelweis (1818–1865) at the Vienna General Hospital in 1848 helped explain the causes of puerperal fever—a deadly condition in postpartum mothers—from a lack of sterilized instrumentation and poor hand hygiene among physicians. Although Semmelweis faced considerable backlash for championing the necessity of rigorous sanitation, including handwashing in a disinfectant solution, an otherwise difficult-to-explain ailment started to be understood, and eventually sparked a revolution in hygiene and sanitary procedure. Two decades later, surgeon Joseph Lister honed these antiseptic methods and ideas surrounding wound healing through his work at the Royal Infirmary in Glasgow.

The 19th century ultimately gave new meaning to human bodies and *normal* human bodies. Increased interest in measurement and averages at midcentury gave rise to concepts like the body mass index while marking the slow decline of individualized humoral and environmental medicine that had dominated previous centuries. In short, the rise of the clinic as *the* location for the production of medical knowledge in the first half of the 19th century allowed for vital conceptual shifts in the meaning of disease which paved the way for the next set of innovations after midcentury.

Despite these radical changes in understanding health and healthcare, clinical medicine also cast

serious doubt on the ability of physicians to cure illness successfully. Observation and organization helped identify disease and harmful practices (like surgery with dirty instruments or excessive use of mercury-based medication) but was less helpful in efforts to find new therapeutics. In some instances, the lack of practical results led to feelings of therapeutic nihilism. Doctors in the 1830s and 1840s questioned their ability to help patients. Despite the rise of medical nihilism, or at least skepticism, by the 1830s, professional identity became increasingly solidified. In the United States, physicians formed the American Medical Association (AMA) in 1847. In previous decades, doctors had formed state and local organizations to share information, provide support, police physician qualifications, and publish medical information. A national organization, however, was something new and emerged at a time when state governments were in the process of eliminating laws governing the American profession.

The systems of the Enlightenment era failed to accurately depict the human body, but they had provided a pathway for therapeutic practice. Clinical medicine could answer questions about what specific ailments were and how they could be identified, but without the physiological or pharmacological context of a larger system the search for cures appeared stuck, or worse, pointless. During that impasse was when chemical and biological laboratories entered the conversation. Nineteenth-century physiology and bacteriology provided an additional scientific context for medicine and new therapeutic avenues that shaped care at the end of the century.

Turn to the Laboratory

By the late 19th century, the focus of Western medicine underwent another major theoretical shift as the laboratory became a site of knowledge production. The tissues-and-lesions focus of some clinical physicians shifted to the level of the cell, microbe, and chemical makeup of bodily fluids. This change was made possible by changes in medicine's allied sciences, especially chemistry. The chemical revolution at the turn of the 19th century expanded European laboratory culture, especially in Germany, and furthered the shift from doctors understanding disease as a disorder

of individual bodies to general ailments governed by the effects of toxins or microbes upon bodies. Although microorganisms had been part of scientific study since the 17th century and the advent of the microscope, the emergence of modern germ theory and association of specific microbes with specific illnesses emerged in the 19th. The search for individual diseases had opened the door for physicians to look for individual and specific causes for those ailments. Advances in laboratory technique—including dyes used to add contrast to cells under microscope examination—borrowed from chemistry, and industry accelerated these changes. Meanwhile, new vaccines could prevent specific infections. Starting with smallpox vaccination at the end of the 18th century, a new class of vaccinations for anthrax and rabies emerged from studies of bodily fluids. Far from the uniform four humors of classical medicine, the liquids that flowed and oozed through human and animal bodies were thought to contain multitudes of substances, some potentially harmful. Medicine entered the laboratory.

The chemical revolution of the 18th and early 19th centuries transformed the understanding of human and animal bodies. While clinical medicine reached its apex in Paris, others looked to chemical analysis and continued research into human physiology as an explanation for health and illness. This was especially the case in the German-speaking universities. From his appointment in Giessen, chemist Justus von Liebig (1803–1873) created the procedures and instrumentation to study the building blocks of living matter: organic chemistry. With this new framework, biochemists saw bodies as living laboratories. Liebig turned scientific attention to living bodies and started to understand life as a process created by different chemical interactions. On this view, illness disrupted those normal chemical process that took place in, around, and across cells and tissues. Meanwhile in Berlin, Johannes Peter Müller (1801–1858) busily trained the next generation of modern physiologists including Theodor Schwann (1810–1882), Hermann von Helmholtz (1821–1894), Karl Friedrich Wilhelm Ludwig (1816–1895), and Rudolf Virchow (1821–1902). These German scientists added chemistry and laboratory experimentation to the search for understanding human life. From those explorations of healthy

life processes, they also formed the theoretical background for a new medicine. Chemical techniques helped explain the function of cells, introduced new staining techniques, and were critical to the growing pharmaceutical industry, which increasingly relied upon synthetic compounds (based on naturally occurring substances) including aspirin, quinine, and opiates.

Alongside research in chemistry, biology laboratories and procedures also led to major changes in the theory and practice of medicine. In the late 19th century, the search for causal explanations of disease and modern germ theory gained new specificity. While the idea that toxins or microbes could cause illness was not new, laboratory technology and improvements to cell staining, incubation, and microscopy meant that the specific connection between pathogen and ailment could be proven analogous to the way in which signs and lesions drove clinical medicine early in the century. In France, Louis Pasteur (1822–1895) leveraged these techniques to create two groundbreaking vaccines, for anthrax in 1881 and rabies in 1885. Meanwhile, in Germany, Robert Koch (1843–1910) and his students became experts in making such connections between pathogens and specific ailments. Koch isolated two key bacteria in the early 1880s that caused tuberculosis and cholera, respectively. Their identification drew medical attention to bacteriological procedure and made the microscope as essential as the stethoscope. Koch's postulates for isolating and identifying pathogens demonstrate the strengths of the new bacteriological era. The steps both emphasize the role of the microbe in disease and the role of the laboratory in defining new diseases. The following steps were designed to allow laboratory scientists to demonstrate that microbes not only caused disease but exactly which microbes were responsible:

1. The microbe can be found in sick/dead organisms but not in healthy ones.
2. The microbe can be extracted, isolated, and grown in culture from the sick organism.
3. The microbe causes the same disease in a previously healthy organism that it did in the original sick/dead organism.
4. The process can be repeated where the same microbe can be isolated from the experimental organism.

Following this procedure, a researcher could, in theory, clearly identify the pathogen responsible for a specific ailment (if it were bacterial). The method also allowed others to replicate and verify results. A generation of *microbe hunters*, young physicians and scientists trained in bacteriology soon operated across the globe often seeking their quarry in colonial settings. Like physicians of previous centuries who participated in colonial projects to learn about new geographies and their diseases, doctors in the late 19th century viewed the *tropics* as spaces home to different microbes capable of infecting the rest of the world. New diseases defined by their causative microorganism continued to be defined in the last years of the century. So too were victor diseases, ailments like yellow fever, malaria, and typhus. Experiments that connect disease transmission to intermediary organisms now had clear chains of transmission rather than being understood as vague ills of the environment.

The laboratory revolution successfully shifted the attention of Western physicians. By the 1880s, more American students were studying in Germany than in Paris, marking the shift from clinical to laboratory medicine on an international scale. These physicians boasted the ability to work with microorganisms and maintain basic laboratory equipment. If the stethoscope and static-electrical machines had characterized earlier eras of 19th-century medicine, the age of the laboratory had its own symbols in microscopes, syringes, vaccines, and antitoxins. Returning to the example of tuberculosis, the disease took on new meaning in this new era as well. The illness known as consumption in the age of systems became the illness associated with tubercular lesions in the clinical era. Only a few decades later, the isolation of the disease-causing bacterium changed the profession's concept of the disease again. Tuberculosis of the late 19th century was a disease caused by a dreaded bacillus that lurked in unsafe bodies, thriving in the crowded and unsanitary conditions of the industrial city.

Therapeutics of course also adjusted to meet the new theories of disease, modern cell theory, and understanding of microorganisms. Researchers such as Paul Ehrlich (1854–1915) in the early 20th century could point to a future in which

anti-microbial therapy was the way forward. Rather than treating bodies and attempting to alter specific physiological processes, medicine would go right to the source and root out the pathogens.

Presenting Alternatives

Not all patients or practitioners adopted the scientific medical trends. The break from 18th-century systems did not result in a clean transition, but rather a fragmentation of medical ideas and practices. The 19th century became an era defined by alternative practices as well as increasingly scientific medicine.

Homeopathic medicine, proposed by Samuel Hahnemann (1755–1843) in 1796, was one of the earliest of the *alternative* systems and has maintained relevance into the present. Like the systems that preceded it, Hahnemann's theories stemmed from a specific relationship between the human body, disease, and nonhuman substances. At its most basic, homeopathy argues that “like cures like.” In other words, a tiny dose of a substance that causes symptoms of illness in a healthy person can be used to counteract corresponding symptoms of illness. Such substances in ultra-diluted form (sometimes up to 10,000 times) formed the basis of a 19th-century homeopath's arsenal. The practice spread from Germany to France and the United Kingdom as well as the United States. In the American context, homeopathy grew into a significant rival for *regular* or *allopathic* physicians, especially among the urban upper and middle classes. Homeopathic physicians regularly found patients and, unlike their allopathic rivals, welcomed women into their schools much earlier.

As the century progressed, other treatments came in and out of fashion, and patients mixed and matched those they found most helpful. In the second third of the century, for example, the water cure gained adherents on both sides of the Atlantic. The United States, however, proved to be an especially fertile ground for alternative medicine. By the 1840s, no state had medical licensing laws (many had been repealed), thereby allowing an open market for practice and practitioners. Alternative systems of medicine and cures not only existed but were able to create schools, and healers could legally use the term *doctor* to describe

themselves. Eclectics, homeopaths, and later in the century, chiropractors, and osteopaths all successfully competed in the open American marketplace. While their theories and therapeutics varied, most maintained some hint of the whole-body systems of previous eras of medicine.

Sarah E. Naramore

See also Cell Theory; Complementary Medicine; Germ Theory; Medicine, Antiquity; Medicine, Contemporary; Medicine, Enlightenment; Medicine, Renaissance

Further Readings

- Baldwin, P. (1999). *Contagion and the state in Europe, 1830-1930*. Cambridge University Press.
- Bishop, P. J. (1980, June). Evolution of the stethoscope. *Journal of the Royal Society of Medicine*, 73, 448–456.
- Burnham, J. C. (2015). *Heath care in America: A history*. Johns Hopkins University Press.
- Clendening, L. (1942). *Source book of medical history*. Dover.
- Gevitz, N. (1988). *Other healers: Unorthodox medicine in America*. Johns Hopkins University Press.
- Hamlin, C. (2014). *More than hot: A short history of fever*. Johns Hopkins University Press.
- King, L. S. (1982). *Medical thinking: A historical preface*. Princeton University Press.
- Owens, D. C. (2017). *Medical bondage: Race, gender, and the origins of American gynecology*. University of Georgia Press.
- Pickstone, J. V. (1992). Dearth, dirt and fever epidemics: Rewiring the history of British “public health,” 1780–1850. In T. Ranger & P. Slack (Eds.), *Epidemics and ideas: Essays on the historical perception of pestilence* (pp. 125–148). Cambridge University Press.
- Porter, R. (2002). *Blood and guts: A short history of medicine*. W.W. Norton.
- Porter, T. M. (2018). *Genetics in the madhouse: The unknown history of human heredity*. University of Chicago Press.
- Rosen, G. (1974). *From medical police to social medicine: Essays on the history of health care*. Science History Publications.
- Rosenberg, C. (1962, 1987). *The cholera years: The United States in 1832, 1849, and 1866*. University of Chicago Press.
- Rosenberg, C. (1979). “Physician versus bacteriologist”: The ideology of science in clinical medicine. In

- M. J. Vogel & C. E. Rosenberg (Eds.), *The therapeutic revolution: Essays in the social history of American medicine*. University of Pennsylvania Press.
- Rothman, S. (1994). *Living in the shadow of death: Tuberculosis and the social experience of illness in American history*. Johns Hopkins University Press.
- Valenčius, C. B. (2002). *The health of the country: How Americans understood themselves and their land*. Basic Books.
- Weiner, D. B., & Sauter, M. J. (2003). The city of Paris and the rise of clinical medicine. *Osiris*, 18, 23–42.
- Weisz, G. (2003). The emergence of medical specialization in the nineteenth century. *Bulletin of the History of Medicine*, 77(3), 536–574.
<https://10.1353/bhm.2003.0150>
- Whooley, O. (2010). Organization formation as epistemic practice: The early epistemological function of the American medical association. *Qualitative Sociology*, 33, 491–511.
- Whorton, J. C. (2002). *Nature cures: The history of alternative medicine in America*. Oxford University Press.

MEDICINE, RENAISSANCE

During the Renaissance, a period in Europe of cultural, political, economic, and artistic *rebirth* that took place from the 14th to the 17th century, there were significant, sometimes dramatic, changes in society and in many fields of endeavor and study, including the theories and practice of medicine. Among the major new discoveries of the era was that of the English physician William Harvey, who published his clinical observations on the circulation of the blood, which overturned the accepted wisdom of the four bodily humors, which had been developed by Hippocrates of Cos (c. 460–c. 370 B.C.E.) and Galen of Pergamon (129–c. 200 C.E.) in ancient times.

Many of these changes and ideas arose from queries about the teachings of the church, and a greater belief in rationalism rather than leaving cures to the power of prayer, which had been argued by St. Benedict and other theologians and had since become orthodoxy. A new emphasis was placed on studying the human body, as is clearly shown in Renaissance art, and which was crucially important for medicine as far greater knowledge of

human anatomy was gained. These new ideas and theories were then able to spread through the advent of movable type and the proliferation of printed books from the 1480s onward, thereby allowing information to circulate rapidly and more freely around Europe. This entry traces the origins of Renaissance medicine in antiquity and the Middle Ages and subsequent growth of anatomical knowledge, continues with an account of the role of universities and research in the continuing development of medicine, and explores the contribution of printing and the growth of literacy in disseminating medical information. The high rates of infection and mortality caused by periodic episodes of the bubonic plague, and the transmission of other infectious diseases and the discovery of immunization are discussed, including diseases introduced by European explorers to the Americas and those brought back to Europe by returning explorers and colonists. The entry concludes with a brief account of progress in medicine outside of Europe before and during the Renaissance, and the exchanges of medical information between cultures previously isolated from one another.

Beginnings

The era of the Renaissance started in Italy about 50 years after the end of the bubonic plague, often called the *Black Death*. The devastation caused by contagious disease killed between 75 million and 200 million people in Europe between 1346 and 1353, from one third to—according to some scholars—up to two thirds of the population of the continent, as well as many people in North Africa and also the Middle East. In Florence, the birthplace of the European Renaissance, the population fell from around 115,000 in 1338 to only 50,000 in 1351. Genoa and Pisa were also devastated, as were many other cities, with Milan being one of the few cities on the Italian peninsula that escaped such devastation. Yet Florence was able to rebound and became wealthy and a center for intellectual thought and ferment.

With many people in Europe feeling, or beginning to recognize, that the resort to prayer was clearly inadequate to deal with problems such as the Black Death, medical science started to concentrate on discovering more about human anatomy. It had long been forbidden for doctors (or

indeed anybody else) to dissect human bodies, with the Greek doctor Aelius Galen and others having worked on animals, especially monkeys, to try to understand anatomy. In ancient times, some of the major advances in medicine had been made in Alexandria, Egypt, where dissections had been allowed and were carried out. Andreas Vesalius (1514–1564), a Renaissance anatomist and physician, wrote that one day when he was out looking for bones on a country highway, he came across a cadaver and used that for his studies. He was to use this and other studies as the basis for his book *On the Structure of the Human Body* which was published in Latin in 1543 and was illustrated with 277 woodcuts, which certainly helped to educate many anatomists and physicians. And certainly during the Renaissance, people such as Leonardo da Vinci and also Michelangelo, working at the Cloister of Santo Spirito in Florence (and indeed many others) were able to study the human body more closely, often being able—in the case of da Vinci—work with the corpses of felons who had been executed. And there was gradually some leeway with dissections. The Dutch artist Rembrandt, in his famous painting *The Anatomy Lesson of Dr. Tulp* (1632), shows how a doctor explained human anatomy to his students; the painting illustrates the opening up of the cadaver's left arm to display the muscles and veins more clearly.

Universities, Medical Research, and Religious Reform

The port city of Salerno in southern Italy had emerged as a major center for the study of religion in medieval times. Some of the early medieval medical experts had been trained there. Gradually the university in the city of Montpellier in France came to rival it, with the University of Paris being heavily supported by the French crown, attracting medical practitioners from England and the Low Countries. And in most major cities in Western Europe, practitioners of medicine were regulated, with barber-surgeons being commonplace. Indeed in 1540 in London, the barbers and surgeons finally came together officially and formed the Mystery and Commonalty of the Barbers and Surgeons of London, a group which later became the Royal College of Surgeons of England.

Salerno, Montpellier, and Paris, however, by no means had the monopoly on medical research. The University of Padua in the north of Italy was also fast emerging as an important center of medical learning. Niccolò Leonico (1428–1524), and the Englishman Thomas Linacre (c. 1460–1524) both trained there. Giambattista da Monte, better known as Montanus (1498–1552), was another physician associated with the University at Padua, where more bedside observations of patients were routinely made by doctors. Theophrastus Bombastus von Hohenheim, better known as Paracelsus (1493–1541) also emerged from the schools in northern Italy. Not all the prominent medical practitioners of the period, however, were trained in the north of Italy. Jean Fernel (1497–1588) who coined the term *physiology* and was the first doctor to describe the spinal canal, came from the University of Paris. Some of the early changes in medical thinking during the Renaissance came from people such as Nicolas Copernicus (1473–1543) of Poland, who was better known as a *physicist* and *astronomer*.

Another astronomer who was to have an effect on medical research was Galileo Galilei of Pisa, who, in 1592, devised the first working thermometer. Hippocrates in ancient times had emphasized the importance of body temperature but there was then no way of measuring this accurately. Galileo's thermometer was able to show changes in temperature, and differences, but it did not have a scale. It was not until 1665 that Christiaan Huygens (1629–1695) devised the centigrade scale to measure temperature, and the German physicist Gabriel Daniel Fahrenheit (1686–1736) devised another scale with made use of much smaller increments. The thermometer was to become extremely useful in studying patients as time progressed and these became more and more accurate, with Anders Celsius (1701–1744) revising the centigrade scale, which also took his name.

The ideas of Copernicus and Galileo challenged some of the major dogmas of the church and represented a new desire for scientific knowledge that took place early in the 16th century, coinciding with the Protestant Reformation, which saw the breaking of the hold of the conservatives within the Roman Catholic Church on medical ideas. The reformers Martin Luther, John Calvin, and others were challenging many of the fundamental teachings of the

church, and with the Bible being translated from Latin into English, French, German, Dutch, and other languages, religious teachings became more accessible to increasing numbers of people and previous interpretations came under question. Although prayer remained important, medical practitioners were able to experiment more and develop their own theories about the causes of ailments and how they might be cured more effectively.

Contagious Diseases, Poisons, and Remedies

The medical profession in the early Renaissance still had to deal with the many diseases and ailments that had afflicted so many people during the Middle Ages, and indeed since ancient times. There were some outbreaks of plague during the early 16th century, with one of the more devastating occurring in 1576–1577 in Venice, where some 50,000 people, around a third of the population, died. Regular outbreaks in London caused the royal court to evacuate the city, and with playhouses, including those associated with William Shakespeare, having to close for fear of the contagion spreading among crowds of people. Leprosy was still common; and smallpox was emerging as a major cause of death. In addition there were the host of common ailments such as influenza and fevers. Malaria continued to pose a major health hazard in some parts of Europe, and the Italians were the first to connect it directly with proximity to swamp lands and *bad air*—hence its name, *malaria*—the connection with mosquito bites not confirmed until the 19th century. One of the main factors that was to reduce malaria during the Renaissance was the need for more agricultural land on account of the large rise in population. This resulted in much marginal land, such as the fenlands in England, being drained, and hence the removal of many of the breeding grounds for mosquitoes.

Poisons were also becoming easier to access, and the prominent Borgia family in Rome became notorious for accusations of poisoning their political rivals and enemies. Food and drink tasters were employed by many noble families, and the Roman Catholic Cardinal John Fisher, the bishop of Rochester, was victim of an attempted poisoning.

His cook, Richard Rice, became, in 1531, the first person in England to be executed by being boiled alive, in an effort to make the punishment a deterrent to anyone else planning to use poison. Antidotes to some poisons were developed, and it was rumored that some Italian political figures daily took a small dose of some poisons to make themselves immune should there be an attempt to kill them. One problem that continued to be faced was also the use of certain poisons in small doses for medical reasons and people accidentally dying when taking, or having been prescribed, the wrong dosage.

Emerging in Europe during this period were newly introduced diseases, of which the most important was syphilis, brought back by sailors in the crew of the ships of Christopher Columbus when they returned from the New World. It spread quickly around the ports of Europe where it became known as the Venetian disease, the French disease, or by any name which indicated that it was passed on by foreigners. Spanish soldiers fighting to defend Naples against the French troops of King Charles VIII spread it among the Neapolitans, and the French soldiers after their occupation of the city in 1495 quickly either succumbed to the disease or took it with them as they retreated through Italy back to France. Gonorrhea also emerged as another important sexually transmitted disease. Its cause, and indeed that of syphilis, was known at the time, with gonorrhea being treated by an injection of mercury. A syringe was recovered by archaeologists from the *Mary Rose*, the Tudor warship that famously sank in the Solent, the straits between the Isle of Wight and the English mainland in 1545, showing the method of injection at the time.

The transmission of disease following the discoveries of Columbus was not all one-way. Many European diseases were taken to the Americas by sailors and conquistadors. Large numbers of indigenous Americans very quickly succumbed to influenza and smallpox—the latter sometimes being deliberately spread. There were documented instances of blankets and other items infected with smallpox or from people who had died of the disease being given to indigenous peoples in North America, showing clear evidence of the Europeans' knowledge of the effects from coming into contact with items from people who had a contagious disease.

The Europeans are also believed to have taken malaria to the Americas, and this was also devastating for many indigenous peoples there. Yet it was also in the Americas that a possible cure for malaria was discovered. It was found that treatment with the cinchona bark, in a powdered form, was able to cure sufferers from malaria. The curative properties of the bark were known from the 1560s and 1570s, but the rarity of the trees led to a massive demand exceeding supply and the treatment becoming very expensive.

It was the Jesuit Brother Agostino Salumbrino (1561–1642), a former apothecary, who found that the Quechua people in modern-day Peru used to chew the bark of a particular tree to relax the muscles and prevent shivering in low temperatures. As shivering was one of the symptoms of some people with malaria, he investigated further. There is a popular tradition that the bark was given to the Countess of Chinchón, the wife of the Luis Jerónimo de Cabrera, Spanish Viceroy of Peru in 1638, curing her, with the story entering into local folklore. It was Salumbrino who labeled the tree as the *fever tree*, and because of his membership in the Society of Jesus, the bark became known as *Jesuit's bark*. It was sold in Europe, and there were many attempts to cultivate cinchona trees in Europe. Superstition was never far from the surface in medical treatment and there were claims that Oliver Cromwell, who grew up near the Fenlands and might have been suffering from malaria for much of his life, refused treatment with the bark because of its association with the Jesuits. King Charles II took the treatment, but in secrecy.

War Injuries and Trauma

Injuries in battle had been commonplace since the first fighting in prehistoric times. However, during the Renaissance, the development of firearms led to changes in surgery, with the French barber surgeon Ambroise Paré (c. 1510–1590) writing his famous treatise in French, *The Method of Treatment for Wounds Caused by Firearms* (1545). A previous writer, Giovanni da Vigo (1460–1525), had suggested that people suffering from bullet injuries should have their wounds treated with hot oil to cleanse the injury. Paré, however, was not convinced by this, and his empirical research on

soldiers after a battle was able to demonstrate that those who were not treated with oil made a far better recovery than those who were. His book established Paré as a pioneer in surgery and on the treatment of battlefield wounds.

There were many religious wars in Europe during the 16th century, and during the first half of the 17th century, much of Europe was devastated by the Thirty Years' War (1618–1648) with England, Scotland, and Ireland badly affected by the Civil Wars of the 1640s. This resulted in large numbers of injuries from fighting, but also a breakdown in law and order, attacks by marauders and deserters, and the spread of diseases. Some saw these wars and devastation as divine judgments on society for its ills. But at the same time, it was a period of much change in medical thinking.

Harvey and the Discovery of Circulation of the Blood

As mentioned earlier in this entry, the most important new theory that emerged was from the English physician William Harvey (1578–1657). Born in England, he had trained at the University of Padua, having become a doctor of medicine in 1602, aged 24. Returning to England, he treated patients at St. Bartholomew's Hospital in London, the oldest hospital in Europe, having originally been founded in 1123 and still occupying its original site. Harvey had been influenced by the work of Michael Servetus (1511–1553), who had studied the blood and the lungs, with Matteo Realdo Colombo (c. 1516–1559) embarking on similar work. Andrea Cesalpino (1519–1603) had actually made references to the circulation of the blood; Harvey, working as court physician to King James I and then King Charles I, later embarked on his own study of blood. This was published in 1628 as *Exercitatio Anatomica de Motu Cordis et Sanguinis in Animalibus* (*An Anatomical Exercise on the Motion of the Heart and Blood in Living Beings*) and his theory, since proven, was about the circulation of the blood around the body, transforming people's understanding of physiology. Initially there was much controversy, with some doctors ignoring his work, and others trying to disprove it as it questioned—and contradicted—the thoughts of Hippocrates and Galen, which had dominated

medical thinking since ancient times. Robert Fludd (1574–1637) and others, however, came to the defense of Harvey, whose work has been acclaimed as being of the greatest importance in the study of physiology.

Further Outbreaks of the Plague

Around the time that Harvey was working on his treatise, there were more outbreaks of the plague. Indeed there had been quite a number during the 16th century and the early 17th century. About a tenth of the population of Amsterdam died in the plague in 1623–1625, and around a million French people died in an epidemic in 1628–1631. The Italian plague of 1629–1631 raged through northern Italy and resulted in the deaths of nearly 300,000 people, especially in Lombardy and around Venice. It was believed to have been spread by German and French soldiers when they took Mantua, and then by Venetian troops, infected with the disease, retreating north. Milan, having been spared the Black Death, lost some 60,000 people, nearly half of the city's population.

The outbreak of the Great Plague in London in 1665 brought back some of the fears of earlier, smaller outbreaks, and some historical comparisons with the Black Death and the devastation of that outbreak more than 300 years earlier, and thoughts of it being divine intervention after 5 years of what Puritans had seen as licentiousness following the restoration of the monarchy in 1660. By 1665, the connection between the plague and rats was noticed, and as a result with the outbreak in 1665, great efforts were made to kill rats, which unfortunately had the effect of the fleas (which actually carried the disease and used the rats as hosts) jumping onto humans and using them as hosts, thereby spreading the disease further. There is also the famous story of the village of Eyam in central England. Some damp clothes had arrived from London and brought with them fleas, with some of the locals quickly succumbing to the plague. Rev William Mompesson of the Church of England, and the Puritan minister Thomas Stanley, both decided to quarantine the village and prevent people leaving and spreading the plague elsewhere. Out of the original population of 350, only 83 survived the 14

months of the plague—although some writers claim that the population of the parish was actually around 800, with just over half surviving. Certainly the villagers were brave and prevented the spread of the disease, but also if they had actually evacuated the village early on, and left it empty for several days, the fleas with no host would have died, and most of the villagers would have survived.

Developments Outside Europe

Outside Europe there were also gradual changes in medical thinking. Smallpox was a major cause of death among Europeans, and this continued to be the case until a smallpox inoculation was developed by the English doctor Edward Jenner in 1796. A study of ancient texts shows, however, that Western medical theories were a long way behind those in Asia, with the possibility that some form of inoculation was being used in India from as early as 1000 B.C.E., and there were certainly inoculations against the disease in China from the end of the 10th century, gradually becoming common by the 16th century.

There was also a division of thought in China between six emerging and different—but by no means indistinct—schools of medical theory. There was still much interest in the complementary principles of Yin and Yang, with doctors feeling that illnesses came from an imbalance between these. Some, known as the *Wen-pou* doctors felt that there was too much emphasis on the *yang*, and that patients should be treated with ginseng or aconite. And there was another school that sought to bring all the presumed five elements and the six vapors into some kind of harmony, which would then cure ailments. A radical group emerged and started to urge for much more drastic medication. This saw men like the physician Chun Szi-sung prescribing arsenic to treat venereal diseases and being remarkably successful. Some conservatives insisted on sticking to the classical medical studies—some of which were several thousand years old. They claimed that there should be no deviations from these works. There was a sixth group, which came to be known as the Eclectic physicians who used a variety of the techniques of the other groups and were adaptable to change.

Acupuncture, developed in China, spread to Korea and Japan, and some of the Jesuit missionaries and others visiting China brought back some knowledge of it with them to Europe. A Dutch surgeon, Jacob de Bondt (1592–1631), described the practice of acupuncture in Japan and Java in his book *De Medicina Indorum*, published in four volumes. In it he describes beriberi for the first time, and also the dysentery epidemic which had broken out on Java in 1628. A colleague, Willem Piso, expanded the work, and a new edition was published in 1658 with Piso contributing a large section on medical treatments in the Americas.

Hermann Buschoff, a Dutch priest working in Batavia (modern-day Jakarta, Indonesia) wrote in 1674 about moxibustion, whereby *moxa* from a dried mugwort was injected into the skin by acupuncture. And in 1682, the Dutch physician Willem ten Rhijne (1647–1700), working in Nagasaki, Japan, described acupuncture for the first time in detail in Europe in his book *Dissertatio de Arthritide: Mantissa Schematica: De Acupunctura: Et Orationes Tres*. He also made a study of leprosy based on his experiences in Asia.

In addition to describing Chinese medical techniques, Western writers also took to China some European medical techniques; many of these were able to be reconciled with existing Chinese herbal treatments and surgical techniques. Herbal treatments continued to be used in China, along with acupuncture. Both were based on detailed records as much as a thousand years old, with a very wide range of treatments being tried and tested. Because of these traditions, there were not that many changes between Chinese medicine and medical techniques from the Middle Ages and those of the Renaissance.

In the Middle East, there were also many similarities between treatment in the medieval period and those used in the Renaissance. Within the Ottoman Empire, Constantinople (now Istanbul, Turkey) was affected by the plagues, as were other cities. This led to various attempts to try to clean up the city, with the plague thought to have arisen from poor living conditions. The link to rats—also established in Europe—was known, but not the connection between fleas and the actual transmission of the plague. However, generally apart from these epidemics, health services around the Ottoman Empire were reasonably good, with

hygiene emphasized by doctors and as a part of the Muslim religion. Holy men helped talk through medical problems with people, administering medicines, as did an increasing number of processional doctors. Herbs were used for some treatments, with the smoking of herbs being recommended to relieve congestion in the chest, with rhubarb used as a laxative, and anise seed to help digestion.

More problematic were psychological illnesses, with some of the treatments involving the pouring of lead. For this, lead was melted in a spoon and then poured into a bowl of cold water. The patient then had to drink some of the water, which was believed to have a therapeutic effect. Another group of doctors were trained in the setting of bones, and used forms and massage in dealing with people suffering from misshapen limbs. Surgery in the Islamic world was a well-practiced art, with surgeons in North Africa and Anatolia known for their skills, which were believed to exceed those of many Europeans. The higher success rate was probably based on the higher hygienic standards that were followed in the Middle East more than in many parts of Europe.

Justin Corfield

See also Medicine, 19th Century; Medicine, Antiquity; Medicine, Contemporary; Medicine, Enlightenment; Medicine, Middle Ages

Further Readings

- Barber, N. (2013). *Renaissance medicine*. Raintree.
- Bell, W. G. (1924). *The great plague in London in 1665*. Dodd, Mead and Company.
- Castiglioni, A. (1934). *The Renaissance of medicine in Italy*. Johns Hopkins Press.
- Chamberlin, E. R. (1965). *Everyday life in Renaissance times*. B. T. Batsford Ltd.
- Cipolla, Carlo M. (1981). *Fighting the plague in seventeenth century Italy*. University of Wisconsin Press.
- Dawson, I. (2005). *Renaissance medicine*. Hodder Wayland.
- Eamon, W. (2010). *The professor of secrets: Mystery, medicine, and alchemy in Renaissance Italy*. National Geographic.
- Gordon, B. L., & Leon, B. (1951). *Medieval and Renaissance medicine*. Philosophical Library.

- Hoeniger, F. D. (1992). *Medicine and Shakespeare in the English Renaissance*. University of Delaware Press.
- Kâhya, E., & Ayşegül, D. E. (1997). *Medicine in the Ottoman Empire, and other scientific developments*. Nobel Medical Publications.
- Kâhya, E., & Ayşegül, D. E. (2003). *Studies of medicine in the Ottoman Empire*. Hamdard Foundation Pakistan.
- Lewis, R. (1971). *Everyday life in Ottoman Turkey*. B. T. Batsford Ltd.
- MacNalty, S. A. S. (1946). *The Renaissance and its influence on English medicine, surgery and public health*. C. Johnson.
- Nutton, V. (2001). *Medicine in the Renaissance city*. Oxford University Press.
- Pagel, W. (1958). *Paracelsus; an introduction to philosophical medicine in the era of the Renaissance*. Karger.
- Park, K. (1985). *Doctors and medicine in Early Renaissance Florence*. Princeton University Press.
- Porter, S. (1999). *The great plague*. Sutton.
- Prinzing, F. (1916). *Epidemics resulting from wars*. Clarendon Press.
- Shefer-Mossensohn, M. (2009). *Ottoman medicine: Healing and medical institutions, 1500–1700*. State University of New York Press.
- Siraisi, N. G. (1990). *Medieval & Early Renaissance Medicine: An introduction to knowledge and practice*. University of Chicago Press.
- Siraisi, N. G. (2007). *History, medicine and the traditions of Renaissance learning*. University of Michigan Press.
- Siraisi, N. G. (2013). *Communities of learned experience: Epistolary medicine in the Renaissance*. Johns Hopkins University Press.
- Tames, R. (1998). *Great plague and fire: London in crisis*. Heinemann.

MEDICINE, 20TH CENTURY

See Medicine, Contemporary; Medicine, 19th Century

MENTAL MODELS

Mental models serve to explain patterns of thought that help people understand their world and the logic for their decisions. A model reproduces the structure of a phenomenon, illustrating the

integration of its crucial features. A scale model might reduce something otherwise too large to study, such as the solar system, or enlarge something too small, such as the DNA helix. A step removed from building a physical model is to represent it using symbolic language for the causal relationships within the structure of the entity. The abstract representation is a conceptual framework that serves as a template for organizing related knowledge. *Model building* refers to an intellectual exercise for constructing such a representation. The following semi-chronological account tracks the evolution of the concept as well as illustrating the connection to whole–part considerations in model building.

Origins and Development

All theories featuring mental models share the legacy of the ancient Greek philosopher Aristotle (385–322 B.C.E.), who noted in the tenth chapter of *Metaphysics* that “The whole is more than the sum of its parts.” Kurt Koffka (1886–1941), one of the German phenomenologists who launched Gestalt psychology, insisted on a more nuanced meaning by saying “whole is something else than the sum of its parts, because summing up is a meaningless procedure, whereas the whole–part relationship is meaningful.” This can be contrasted with behaviorist studies of discrete psychological phenomena: The shift was one of perspective and scale from the molecular to the contextual, from the mechanistic to the organic, from inequitable differences to meaningful commonalities, or as Buckminster Fuller labeled it, *synergy*. The perspective changes from a two-dimensional cause and effect sequence to a three-dimensional structure of multiple simultaneous influences. There is a further emphasis on the active and proactive compared to the passive and reactive. This family of theories is therefore dynamic, attending to the interactions and responsive processes in what are seen as functioning systems. The economist and political theorist Karl Marx (1818–1883) introduced the term *interdependence* in the Communist Manifesto as part of the argument against social alienation and to emphasize the importance of establishing cooperation among autonomous individuals. Kurt Lewin (1890–1947) formulated an equation to represent this contextual influence: $B = f(P, E)$, that is,

Behavior is a Function of the Person in the Environment, as part of his *field theory* that required a present context in order to understand the individual. Lewin is also credited with the term *group dynamics*, which remains central to much organizational theory, as does Morton Deutsch's model of *conflict resolution*. Urie Bronfenbrenner (1917–2005) developed an *ecological systems theory* incorporating five environmental systems, expanding outward from the individual's immediate sensory context through micro- to macrosystems.

Principles of Mental Models Theory

An important principle in mental models theory is that each individual will construct meaningful models for understanding their world and that the individual mental models will always feature some variation between people. Further inconsistency may be noted within the individual who will have developed innumerable such models for concepts. Contrary information creates *cognitive dissonance*, upsetting the equilibrium described by the old mental model. This becomes a problem to be solved, often with a domino effect because the component parts are interdependent and dynamic. Mental models theory holds that individuals are continually constructing and revising these cognitive maps. Once formed, the mental models explain the persistence of false beliefs due to a preference for a familiar orientation; that is, a reliance on what one has assumed is a reliable map. Aligning one's unique set of mental models is the hallmark of maturity and integrity, while a fragmented capacity to understand or act is associated with mental instability if not insanity. Philosophically, this might be termed an existential crisis, revealing the close alignment of mental models theory with existentialism; in other words, the focus of understanding reality in terms of one's experienced existence rather than some preordained destiny or divine purpose. Mental models as a term is thus chiefly associated with cognitive science; that is, psychologies of perception, representation, and learning.

There are multiple theories to explain the process of finding a logically consistent balance among many separate concepts. First, according to the Swiss developmental psychologist Jean Piaget (1896–1980), *equilibration* is a biological necessity; in other words, the internal mental structures

must be aligned and realigned to maintain balance amid multiple, overlapping concepts. Piaget proposed two possible responses: assimilation, or the addition of new information into an existing framework, and accommodation, or the adjustment of the framework to align with the new information. Second, according to *cybernetics* theory, feedback occurs, which prompts adjustment in one's mental model. Third, Phillip Johnson-Laird proposed a unified theory of reasoning that uses mental models to integrate three patterns of thinking, namely the inductive, the deductive, and the probabilistic—and their many products of understanding, such as inductions, explanations, estimates of probabilities, and informal algorithms. He explained that the mind represents each possibility in a separate mental model and then is able to infer what remains constant among all of them.

Executive Functions, Worldviews, and Statistics

The ability to analyze a complex concept in order to extract its essential components and functions is considered a higher level of thinking than to merely recall information, as is the ability to create something representing the cohesive whole. These *executive functions* are used to understand the world and to construct internal mental structures that can be further generalized to novel scenarios. This process may seem intuitive if prompted by extensive experience that allows patterns to become predictable, but it is nonetheless hard work. For this reason, once an individual has formed a mental model, it becomes a reliable tool for further understanding, and one is reluctant to *reinvent the wheel*, or in colloquial language, “Don't fix what ain't broke.” It is therefore reasonable that false beliefs persist in the face of contrary information because so much effort has already been invested in the incumbent conceptualization. Studies in adult development have found the increasing tendency to reduce details to gross generalizations, especially after age 55, an important factor to consider social culture as the human population lives longer. As technology has increased along with competition for attention, mental models theory is helpful for explaining the difficulty of reasoning while monitoring stimuli. Mental models are useful for researchers to study which distractions most impair higher level

reasoning. Because verbal stimulation appears to interrupt reasoning to a greater degree than simple rhythmic auditory stimulation, there is more support for the idea that mental models do not merely represent an understanding of reality in symbolic form but are actually the language of understanding. The noted linguist Noam Chomsky has argued that language is not just a tool but the actual mechanism of thought. Computer languages try to reconstruct mental models in order to simulate human reasoning.

Many names—for example schema, conceptual framework, belief system, worldview, mindset, theoretical orientation, and even philosophy—have been used interchangeably to suggest an integration of related ideas into a cohesive and functioning whole that represents reality. The *individual psychology* of the Austrian psychoanalyst Alfred Adler (1870–1937) suggested that people develop worldviews that serve as templates for finding sense in a random world. Martin Seligman promoted *positive psychology*, that is, the study of human potential instead of the abnormal and pathological, which addressed three types of happiness: (1) pleasure and gratification, (2) embodiment of strengths and virtues, and (3) meaning and purpose. While the first may be seen as an emotional response and the second a matter of identity and self-efficacy, the third suggests a conceptual framework about life wherein one influences one's own perspective. Thus mental models are a useful tool for appreciating the complexity of living in a complex world. The sophistication of modern statistical packages is due to the capacity to calculate correspondence among many variables; the quality of a study rests on its capacity to fully describe the relationships among them and to identify rival hypothesis and confounding influences that complicate the usefulness of the findings.

Mental Models in Communication and Teaching

Efforts to organize ideas are not new. The great rhetorical traditions feature coherent, logical arguments; a disorganized rant is less likely to persuade. Master communicators, whether in the classroom or the pulpit or on a campaign trail, have long used figurative language to help explain concepts and inspire action. By use of analogy or

metaphor, one is led to transfer an understanding of some aspect of a familiar phenomena to a novel phenomena. The underlying structure is thus abstracted and a sense of truth may be developed to predict and explain reality. To use a travel metaphor, the mind can more easily navigate the crowded landscape of phenomena by using a cognitive map for orientation. Poetry is powerful in its capacity to envelop the listener in a representation of combined sensory perception. Aphorisms, parables, and allegorical stories become clichés within a cultural literacy because of their enduring value to make efficient sense out of the counterintuitive or confusing.

We are witnessing shared mental models emerge in response to news as varied as the abduction of Nigerian school girls (#bringbackourgirls) or the fear of climate change or the use of Common Core math standards. Finding out others' mental models helps predict compatibility of purpose and method, and thus mental models are a useful means to discover the degree of alignment among a group of people who must collaborate. Mental models are thus useful at a corporate level as well. Schools of thought are more advanced articulations of shared beliefs. These conceptual frameworks, for example, political ideologies, fitness regimes, pedagogical methods, religious doctrines, military strategies, and artistic genres, in turn provide a focus for affiliations as well as feuds. Many debates reveal another dimension, the mental model of corporate effectiveness. Team performance has been found to correlate with *team mental models* (TMMs). Teams in any domain, such as sports or business, work to achieve outcomes requiring complex decisions. The profiles of shared cognitions are further distinguished between common knowledge that varies little between teams and team-specific knowledge. The principle of *equifinality* explains why the specific knowledge is not as significant as the degree to which that knowledge is shared: Several different TMMs in the same domain may be effective. Thus, the production of intended outcomes is informed by measures of agreement. This recent line of organizational research relies on increasingly sophisticated multivariate statistical analyses to describe the agreement within teams as well as between teams, resulting in new formulas to reduce data and indices to summarize shared cognitions. The objective

of establishing team cohesiveness serves the ultimate goal of team effectiveness. Similarly, socializing individuals to function collaboratively is an important part of school curricula because research has found that people who lose their jobs suffer not from ignorance or incompetence so much as incompatibility. Disagreements often hinge on dispositions; that is, the profile of beliefs, attitudes, and values that justify decisions made and roles played. Organizations therefore try to inspire loyalty and consistent performance by providing guidelines in the form of mission statements, mantras, strategies, and dogmas that are memorable for their simplicity. Logos and branding are key, but they serve a more nuanced purpose if they symbolize the structure of the model instead of merely providing a distinctive motif.

The theory of mental models is revealed in the use of visual strategies to improve teaching and learning. As a tool, mental models are used to influence comprehension for a great range of purposes, such as teaching the causes of a war or the function of the circulatory system, or for promoting strategic communication and decisive action. *Advance organizers* present mental models that highlight crucial relationships of elements before the component parts are understood fully. Extensive educational research supports their use and in turn inspires the development of more theories of learning. The reverse process can be seen in a *concept map*, which is constructed in order to show the relation of component parts to the main concept. This can be used as an assessment tool to determine whether a concept has been learned, or as a brainstorming tool to help a team collect all significant issues for consideration related to the team's goal. A concept map may also be used as a writing tool to promote cohesive thought, often called a *word web*.

Mental models are compatible with many practical applications in education. Gardner's multiple intelligence theory posits that a concept can be understood in several ways. From a mental model perspective, multiple intelligences can be seen as different ways in which to perceive and organize one's understanding. The use of manipulatives to learn mathematical concepts is now standard practice. Teacher candidates are cautioned to consider the great range of student differences and thus expected to use a variety of methods in response to student characteristics. Assessment instruments are

no longer limited to selected response items with one right answer. Rather, teachers are expected to use constructed response items, such as reflective essays or projective drawings or concept maps, to find out what and how well students have learned the target concepts. Teachers are also cautioned to consider the cultural influence of the models they use, for instance, the Mercator projection of the three-dimensional globe to a two-dimensional map distorts the relative size of the dominant, Western areas, which in turn may be regarded as a disservice to underrepresented societies. Finally, schools are now held responsible for interrupting students' negative mental models of exclusion and privilege, in part because antisocial behavior is influenced by prejudice. Thus, when teachers are interviewed, they are asked to provide a statement of their beliefs about teaching and learning; that is, they must articulate their own mental model of their role in the profession.

Naomi Jeffery Petersen

See also Abstraction; Belief Revision; Framework; Understanding; Worldview

Further Readings

- Chowdhury, P. D., Haque, C. E., & Driedger, S. M. (2012). Public versus expert knowledge and perception of climate change-induced heat wave risk: A modified mental model approach. *Journal of Risk Research*, 15(2), 149–168.
- Gary, M. S., Wood, R. E., & Pillinger, T. (2012). Enhancing mental models, analogical transfer, and performance in strategic decision making. *Strategic Management Journal*, 33(11), 1229–1246.
- Gogus, A. (2013). Evaluating mental models in mathematics: A comparison of methods. *Educational Technology Research & Development*, 61(2), 171–195.
- Johnson-Laird, P. N. (1983). *Mental models: Towards a cognitive science of language, inference, and consciousness*. Cambridge University Press.
- Mumford, M. D., Hester, K. I., Robledo, I. C., Peterson, D. R., Day, E. A., Hougen, D. F., & Barrett, J. D. (2012). Mental models and creative problem-solving: The relationship of objective and subjective model attributes. *Creativity Research Journal*, 24 (4), 311–330.
- Ragni, M., & Knauff, M. (2013). A theory and a computational model of spatial reasoning with

- preferred mental models. *Psychological Review*, 120(3), 561–588.
- Rook, L. (2013). Mental models: A robust definition. *Learning Organization*, 20(1), 38–47.
- Schaffernicht, M., & Groesser, S. N. (2011). A comprehensive method for comparing mental models of dynamic systems. *European Journal of Operational Research*, 210(1), 57–67.
- Sinreich, D., Gopher, D., Ben-Barak, S., Marmor, Y., & Rakefet, L. (2005). Mental models as a practical tool in the engineer's toolbox. *International Journal of Production Research*, 43(14), 2977–2996.

METAPHYSICS

Metaphysics is the study of the fundamental nature of reality. It is concerned with questions of existence and its nature. Ontology, the study of what exists, is a branch of metaphysics. The subject is extremely broad and as old as philosophy itself, so a comprehensive survey of the subject and its history is beyond the scope of this entry. Instead, following some brief background remarks, the entry will focus on giving a detailed example of one traditional metaphysical debate that serves both to illustrate the types of concern that metaphysicians deal with, and to give a sense of how the subject's history connects with its modern form. Following some further remarks about the fall of metaphysics in the 17th century and its revival in the latter half of the 20th century, this essay concludes with a brief overview of the influence of metaphysics in other areas of philosophical inquiry and in science.

Background

Although Plato and the Presocratics present metaphysical arguments, the first work to bear the title *metaphysics* is a selection of writings by Aristotle, written in the 4th century B.C.E. Aristotle described the subject matter of these writings as *first philosophy* but did not himself use the term *metaphysics*. This title was bestowed by Andronicus of Rhodes, the editor of Aristotle's works who collected the writings together in the 1st century C.E. At that time a significant part of one's philosophical education would have consisted in a

close reading Aristotle's works, and that Andronicus called the selection *metaphysics* reflects the fact that the writings contained therein were to be studied after (i.e., temporally) Aristotle's work on nature (*ta phusika*). So, although the etymology of the term *metaphysics* suggests that its subject matter is *that which goes beyond physics*, this is a misnomer. Even knowing this, however, one might still be puzzled about the distinction between metaphysics and physics. The reason is that one might reasonably describe physicists themselves as enquiring into the fundamental nature of reality. The methods employed by physicists differ markedly from those employed by metaphysicians, however, and they deal with questions at different levels of abstraction. Physicists are primarily concerned with the existence of particular physical particles, properties, and laws. Metaphysicians, by contrast, are more concerned with what kinds of entities physical particles, properties, and laws *are*, and with whether any of them is (or are) more fundamental than the rest. And while physicists primarily employ empirical methods to establish their claims, metaphysicians primarily employ *a priori* methods to establish theirs. However, as the last section of this entry will outline, the relationship between metaphysics and physics (and science more generally) is nonetheless a more intimate one than this might suggest.

Examples of questions that metaphysicians have sought to answer include (1) What is the relationship between the mind/mental properties and the body/physical properties? Are they distinct, or identical? (2) What is time? Does it flow? Do the past and future exist as the present does? Is time travel possible? (3) How is persistence over time possible? How can one and the same thing have different properties at different times? (4) Do we have free will? Is determinism compatible with our having it? (5) What is causation? Do causes necessitate their effects? (6) Is there a God? If so, what is His nature? (7) Could there have been nothing at all? More generally, which events are possible, and which impossible, and what makes them so? (8) Do abstract objects such as numbers exist? Are mathematical truths mind-dependent or mind-independent? (9) Do the fundamental entities posited by science, such as electrons and quarks, exist? Or are they just useful fictions? (10) What kinds of entities are objects and properties,

and what is the relation between each object and the properties it possesses? Is either more fundamental than the other?

Each set of questions has generated a substantial literature (both historical and contemporary) that continues to grow, and so it is not possible to cover each in detail here. In lieu of that, in order to give the reader a better grasp of the kinds of issues that are dealt with in contemporary metaphysics, and how they connect with the history of the subject, the last set of questions is examined in more detail in the next section. (Some of the other questions are discussed in other entries in this encyclopedia, and interested readers are encouraged to consult the cross-references in addition to the further readings for more details.)

Objects and Properties

The term *object* (or *substance*, or *particular*) is used by metaphysicians to pick out individual things that bear properties, but that are not themselves properties (this latter qualification is needed because properties can themselves have higher order properties). Everyday examples of objects include people, cars, tables, and houses. But larger physical items such as planets and stars, and smaller physical items such as atoms and electrons, are also objects. In addition, abstract entities like numbers (if such there be) are objects. The term *property* (or *quality*, or *attribute*) is used to pick out ways that objects (and properties themselves) can be. Everyday examples of properties are those possessed by everyday objects. People can have the property of *being tall*, cars of *being fast*, tables of *being brown*, and houses of *being semidetached*. But all objects, it seems, bear properties of some sort. Every star and atom, for example, has a certain mass, every electron is negatively charged, and every (natural) number is either odd or even. This raises the question: What is the relationship between an object and its properties?

Realists maintain that objects and properties are distinct but equally fundamental kinds of entity. An object possesses a property, on this view, by bearing some kind of relation to it. An important motivation for *realism* is the felt need to explain how two objects can be the same in certain respects. Consider: This apple is red, and

so is that table, so the apple and the table are identical in that respect. An appeal to the existence of properties is supposed to offer a way to explain how this can be. Views of this kind predominated from ancient times until the early medieval period, when *nominalism* (discussed in the next paragraph) began to gain ground. Objects are generally viewed by realists as being unitary nonrepeatable entities that exist by having parts that occupy single spatiotemporal regions. But there is less agreement regarding properties. The two main traditional realist conceptions of properties are those of Plato and Aristotle respectively. Plato's view is that properties are immutable entities called *Forms* that exist in a supersensible realm. Objects possess properties on this view by standing in a special *participation* relation to them. Aristotle's view (on the most common interpretation) is that properties are entities called *universals* that exist wholly and immanently in each object that possesses them. The Platonic view came under heavy criticism in the medieval period and had all but disappeared before the Renaissance, but Aristotle's view has proved more durable, and has survived until the present day. Currently, the most prominent modern defender of the Aristotelian view is perhaps the Australian philosopher David Armstrong. On Armstrong's version of realism, universals are nonspatial parts of the objects that possesses them. This account has the advantage of being able to explain in a particularly simple way how two objects can be the same in certain respects; they can be so simply by having a numerically identical nonspatial part in common. It is worth noting, however, that on Armstrong's view it is only those universals that correspond to fundamental physical properties that exist, and all other resemblances are explained in terms of the sharing of these universals.

In contrast to realists, nominalists hold that objects are the only fundamental entities that there are. They either deny the (mind-independent) existence of properties altogether, or explain how their existence is parasitic upon the existence of objects. The 4th-century Roman philosopher Boethius was the first to discuss this view in any detail, but it was in the period between the 11th and 16th centuries that the view gained widespread support. Important defenders of the view

in this period include Peter Abelard, Thomas Aquinas, and William of Ockham. The latter's view is that properties exist only in the mind as concepts that apply to many things at once. Although similar nominalist views have been expressed by major historical philosophers since this period, Ockham's work serves as the pinnacle of historical thought on the position. Only in recent times has it become clear that nominalist positions are available that allow properties to be mind-independent (nonfundamental) entities. According to one of the most popular views of this kind, properties are nothing but sets of objects. So, for example, the property of *being brown* is just the set of all objects (in this case physical objects) that are brown. This view is most closely associated with the 20th-century American philosophers Willard van Orme Quine and David Lewis. Both reject the contention that we must admit the existence of properties so as to explain how two things can be the same in certain respects. Instead they maintain that facts of this sort are basic facts that require no further explanation. This, combined with their adherence to a methodological principle known as *Ockham's razor* (viz. that we should avoid a commitment to additional kinds of entity wherever possible), is a major motivation for their nominalist position.

Bundle theory, a view that has only come to prominence in recent times, reverses the order of dependence invoked by nominalists. According to bundle theorists, properties are the fundamental entities, and objects are nothing more than com-present collections of them, and so parasitic upon them for their existence. This progenitor of this view is the 18th-century Scottish philosopher David Hume who held that the self is nothing but a bundle of perceptions. Those who adopt bundle theory are generally motivated by an apparent epistemological difficulty that surrounds the concept of an object: If an object is a bearer of properties but not itself a property, then it is difficult to see how we can form any positive conception of what an object is, for if we try to conceive of any object in abstraction from its properties, there is nothing remaining for us to form a conception of. For those with empiricist leanings, this consideration casts considerable doubt on the coherence of the concept of an object. Indeed, it was John Locke, a committed empiricist, who first noted the

difficulty in the 17th century. Locke suggests, though, that despite this difficulty, we must retain a commitment to objects in our ontology in order to explain what ties together properties into the distinct clusters of them that we face in experience. He suggests, in other words, that objects are needed to serve as the things in which properties inhere. Bundle theorists, however, believe that they can explain everything that needs explaining here, while avoiding a commitment to objects altogether, by maintaining that properties can be bound together into bundles by standing in a mutual relation of compresence.

Bundle theory is, at least *prima facie*, compatible with the view that properties are Aristotelian universals. But despite there being some notable defenders of just such a view (e.g., Bertrand Russell in the 1940s) most have been convinced that it is untenable by the following argument: (1) If objects are nothing but bundles of Aristotelian universals, then it is impossible to have two exactly similar objects (for *they* would be bundles of numerically identical universals, and so *they* would in fact be one and the same bundle). (2) But it is possible to have two exactly similar objects. C. So, objects are not bundles of Aristotelian universals. This argument was put forward forcefully by the British-American philosopher Max Black in 1952. In the following year another American philosopher, Donald Williams, published a highly influential defense of bundle theory that avoids this problem by conceiving of properties as being, like objects, unitary nonrepeatable entities that are identical with their instances. Williams coined the term *trope* for entities of this sort, and since then the literature on bundle theory has been dominated by the trope variant of the view. Many of the reasons put forward for believing that properties are tropes, however, do not involve a commitment to bundle theory, and many who reject the view have found a place for tropes alongside objects, instead of universals, in their ontology of fundamental entities.

The Fall and Rise of Metaphysics

The foregoing description of the debate surrounding objects and properties is necessarily far from exhaustive. For example, it makes no mention of those who admit further fundamental entities into

their ontology (e.g., Jonathan Lowe who holds that there are four kinds of fundamental entity). Nor does it mention those who reject the object/property dichotomy altogether (e.g., Alfred North Whitehead, who holds that the fundamental entities are neither objects nor properties, but events). Nonetheless, it should furnish the reader with a good understanding of what metaphysics is. As the historical details given in the description suggest, the debate about the relation between objects and properties died down in the 17th century and only came to prominence again in the latter half of the 20th century. As it was for this debate, so it was for traditional metaphysical debate in general. This is not to say that important work in metaphysics did not take place during this period of respite, but as epistemology took center stage in philosophy during the 18th and 19th centuries, a great deal of traditional metaphysics was sidelined, and in fact directly attacked by significant philosophical figures (e.g., by Hume and Immanuel Kant). This negative attitude was strengthened during the first half of the 20th century under the influence of two major philosophical movements: *logical positivism* in the 1930s and 1940s, and *ordinary language philosophy* in the 1940s and 1950s. Those from each movement were apt to complain that traditional metaphysical claims are either meaningless or confused. As these movements died away—and owing to the work of a number of pioneering philosophers in the 1940s and 1950s who were unafraid to tackle metaphysical issues head-on despite the prevailing negative attitude—metaphysics slowly regained its status as one of the central areas of philosophy. Quine, mentioned earlier, deserves special emphasis as being particularly influential in this regard. In the 1940s, he expounded a criterion of ontological commitment utilizing the methods of formal logic. According to this criterion, one gains a commitment to particular objects by accepting a given discourse (e.g., to numbers by accepting mathematical discourse). Although Quine's criterion has not been universally accepted, discussion and development of Quine's work fueled the revival of metaphysics. By the mid-1980s the revival was complete and is still going strong today.

The Influence of Metaphysics

As a foundational area of inquiry, the influence of metaphysics in other areas of philosophy is profound. Indeed, the influence in certain other areas is so extensive that those areas are best thought of as being subareas of metaphysics itself. For example, metaphysics predominates in the philosophy of mind, where the major questions regard the metaphysical nature of the mind/mental properties and their relation to physical objects/properties. Similarly, in the realm of personal identity the major questions regard the metaphysical nature of persons over time. But even in areas of philosophy where metaphysical concerns are not primary, metaphysics plays a central organizing role. This is true in ethics, epistemology, aesthetics, the philosophy of science, and practically all other areas. Again, a full survey of its influence is not possible in this short entry; instead an illustrative example is provided.

In ethics, one fundamental question concerns the metaphysical nature of moral properties, and seminal contributions in this area have focused on just this issue. Historically, for example, Plato conceived of goodness as a Form (i.e., the Form of the Good), knowledge of which is necessary for any other knowledge to be of benefit to us. In more recent times, G. E. Moore has advanced the position (in *Principia Ethica*) that moral properties are *sui generis* (of their own special metaphysical kind) non-natural properties that we come to know via a special faculty of moral intuition. This position, though now widely rejected (due in large part to John Mackie's argument that such properties are "metaphysically queer"), has served as a touchstone in contemporary debate, with many developing the position that moral facts are natural facts, a view known as *metaphysical naturalism*. Such positions have been developed by Michael Smith and Phillipa Foot, among others.

The influence of metaphysics in the sciences is a more contentious matter. Some preeminent scientists have gone so far as to claim that philosophy as a whole has no bearing whatsoever on science. Stephen Hawking, for example, famously declared that philosophy is *dead* on the grounds that philosophers have not kept up with developments in science. However, Hawking's claim provoked a robust response from other scientists and philosophers, with many suggesting that Hawking

himself had failed to keep up with developments in philosophy, and in particular metaphysics. Critics of Hawking's view in both science and philosophy have pointed out that the interpretation of scientific theories is itself a metaphysical enterprise, and that philosophers and scientists work fruitfully alongside each other in an effort to clarify and understand our best scientific theories. Consider the questions: (1) Is space-time in general relativity best thought of as an independent substance? (2) Does the mind play a fundamental role in the collapse of the wave-function in quantum mechanics, as currently understood? (3) Can neuroscience give us anything more than correlations between brain states and states of consciousness? Each of these illustrative questions, properly understood, have both scientific and metaphysical elements. Whereas scientists are experts in the empirical and mathematical methods necessary to obtain a clear view of how to answer them, metaphysicians are experts in the conceptual methods that are no less necessary. To understand the questions, we need to know the empirical facts, which is what science provides; but we also need to know how best to conceptualize the structure of those facts, which is what metaphysics provides. As such, if we want to progress beyond a purely instrumental understanding of science, metaphysics and science must be considered as integrated disciplines with the same fundamental goal: that of understanding the fundamental nature of reality.

Benjamin L. Curtis

See also Constructive Empiricism; Instrumentalism; Intuitionism in Logic and Mathematics; Phenomenalism; Philosophy of Mind; Philosophy of Science; Pragmatism; Realism in Mathematics; Scientific Realism

Further Readings

- Oliver, A. (1996). The metaphysics of properties. *Mind*, 105(417), 1–80.
- Rea, M. C. (Ed.). (2009). *Arguing about metaphysics*. Routledge.
- Russell, B. (1946). *History of Western philosophy*. G. Allen and Unwin Ltd.
- van Orman Quine, W. (1948). On what there is. *Review of Metaphysics*, 2(5), 21–36.

METATHEORY

A metatheory is a set of beliefs, principles, and assumptions that people hold about what is real (ontology) and how individuals come to know reality (epistemology). Different metatheories involve different sets of beliefs and assumptions, but some form of metatheory underlies and conditions all theoretical and empirical work undertaken in the sciences, engineering, and mathematics. As sets of background concepts and presuppositions, metatheories demarcate what is sensible, meaningful, and appropriate in the theories and methods constructed to explain the world. Whereas theories take empirical phenomena (i.e., things and activities in the natural world) as their objects of study, metatheory takes theories themselves (and concepts more broadly) as the objects of study. As a practice, therefore, metatheorizing involves philosophical inquiry into—in other words, conceptual analysis of—theory and method. Metatheorizing encompasses a diverse cross-section of activities. These include philosophical analysis and synthesis of the background concepts and assumptions of metatheories themselves, as well as examination of what metatheoretical frameworks underlie and inform different theories and methods of a discipline. They also include application of specific metatheoretical principles, such as beliefs and assumptions about what constitutes good science or theorizing, to the analysis of theory and method itself.

Mapping the Conceptual Space of Metatheory

Basic *common sense* or *folk* beliefs and assumptions about the nature of reality—about how the world is organized, for example, or what being alive and being human entail—inform everyday human experience, constraining how people perceive and conceive of the world and themselves. These various beliefs and assumptions develop as a function of people's lived engagement with their physical and cultural environments. However, they remain largely backgrounded in the sense of being implicit and not something of which most people are consciously aware. A critical form of metatheorizing occurs when philosophers render explicit these commonsense beliefs and assumptions and

then proceed to organize and systematize them into coherent, principled sets, namely metatheories.

Stephen C. Pepper's book *World Hypotheses: A Study in Evidence* provides a paradigmatic example of this kind of metatheoretical activity. In his classic and heavily influential 1942 publication, Pepper argued for a total of six different metatheoretical frameworks of belief and assumption, only four of which he considered to be adequate models of reality: *formism*, *mechanism*, *contextualism*, and *organicism*. Each framework, or worldview, offers a distinct characterization of the nature of reality and truth, represented by a defining root metaphor. For example, likening the universe and everything in it to a machine constitutes the root metaphor of a mechanistic world view, whereas likening the universe and everything in it to an organic whole constitutes the root metaphor of an organicist world view.

Continued conceptual analysis of Pepper's world view designations has established possibilities for further syntheses. Willis F. Overton, for example, has argued that properties of a contextualist world view, such as its characterization of the universe as inherently active and ever-changing, can be subsumed under the integrative, holistic focus of organicism to yield a synthesis of both worldviews. This overarching worldview framework, which Overton has termed the *process-relational paradigm*, emphasizes multiple, irreducible levels of organization in reality and highlights the importance of adopting multiple epistemological perspectives on any given phenomenon. Its metatheoretical rival, termed the *Cartesian-split-mechanistic paradigm*, is akin to Pepper's mechanistic worldview and emphasizes the reduction of all complexity in the universe to one true reality: a foundational set of fixed elements or substances (e.g., atoms) of the universe. These elements exist independently of one another, and, through linear associations with one another, serve as the driving force behind the bottom-up generation of all structural complexity.

Overton's contrast between process-relational and Cartesian-split-mechanistic paradigms shares many similarities with another prominent worldview divide, namely between process and substance ontologies. Whereas a *process ontology* considers processes, rather than substances or things-in-themselves, as primary to all existence, *substance ontologies* characterize the universe's foundational

state in terms of fixed, elemental substances. Some form of substance ontology has conditioned much of scientific thought since the 17th century. Since the end of the 20th century, however, increasing calls of support for the adoption of its metatheoretical rival, process ontology, have emerged across a variety of scientific disciplines, as illustrated, for example, in Daniel J. Nicholson and John Dupré's 2018 edited volume *Everything Flows: Towards a Processual Philosophy of Biology*.

Uncovering the Metatheory Behind Theory and Method

Metatheorizing about metatheory is an exercise largely confined to the realm of philosophical discourse. Another form of metatheorizing, however, extends conceptual analysis to actual theory and method in scientific and other disciplines. In this form of conceptual analysis, various theories and methods are subject to critical examination in order to establish the metatheoretical beliefs and assumptions upon which they are founded. Recall that such underlying beliefs and assumptions profoundly affect how scientists think about their respective disciplines, from defining what constitutes a problem to be investigated and what questions to ask, to how to engage in empirical work or interpret empirical findings. As a result, explicitly identifying the metatheory that informs and constrains empirical and theoretical work offers a critical guide for understanding the conduct of all theoretical and methodological activity.

For an example of how this form of metatheorizing can unfold, consider the classic nature versus nurture debate in the psychological and life sciences. As typically conceived, the debate concerns whether nature (in the form of endogenous factors such as genes) or nurture (in the form of exogenous factors such as the physical and cultural environment) is primarily responsible for the development of an organism's traits. Both theoretical camps involved in this debate would seem to have little in common. However, conceptual analysis reveals some striking similarities in their background assumptions. Both nature and nurture theories model organisms as passive, as mechanistically acted upon and driven either by forces outside them or by forces within them. Both also presuppose that *nature* and *nurture* constitute independent sources

of trait development—sources that may combine additively. Both theories, in other words, derive from a common worldview: mechanism. This has enabled *resolution* of the debate in the form of integrated theories that treat both nature and nurture as interacting sources of development.

Debate still persists, however, over precisely what is entailed by forces of nature *interacting* with forces of nurture. At this level of the debate, conceptual analysis reveals a deeper, metatheoretical divide over how to conceptualize the interactions that occur between nature and nurture. On one side of this metatheoretical divide are those who see interactions as decomposable into distinct nature and nurture components, each of which plays a distinct role in the formative process. For example, factors of nature predispose toward a certain trait, but factors of nurture are needed to trigger, or actualize, that predisposition. From this metatheoretical vantage point, factors of nature and factors of nurture are ontologically independent and can therefore be legitimately separated through methodological analysis to determine relative levels of influence. On the other side of this metatheoretical divide are those who see the dynamic process of interaction itself as the driver and source of all traits. For example, an organism's traits constitute genuinely emergent products of interaction, irreducible to the *components* of the interaction themselves. From this metatheoretical vantage point, it makes no sense to think about nature and nurture as separable or different sources of traits. Likewise, methodologies that seek to study the distinct influences of nature relative to nurture are considered inherently faulty.

This extended example illustrates how important it is to reveal what metatheoretical frameworks underlie theory and method. When conceptual analysis indicates that two competing theories share common metatheoretical assumptions, then such theoretical debate can be readily resolved through empirical means. However, if a theoretical debate reflects a metatheoretical divide, then that debate exists beyond the realm of empirical investigation and therefore cannot be adjudicated through empirical activity. Such debates are philosophical in nature because different metatheories operate according to different truth criteria concerning the nature of reality and what methods are appropriate for knowing that reality.

What Makes Good Science: Evaluating Theory Using Metatheory

Metatheoretical activity is crucial for the development of a healthy scientific discipline, yet such activity also takes place outside the realm of science. In mathematics and logic, for example, metatheories often function as contexts in which theories and operations can be examined, as if from the outside. A metatheoretical level includes all operations of the level *beneath* it and therefore gives mathematicians and logicians an objective perspective from which to analyze the characteristics and functions of those operations and to formalize metatheoretical operations that subsume them. Likewise, in linguistics, a metalanguage is a language that is built to analyze the elements of an *object* language without being limited by, or even contained within, that language.

In scientific disciplines, metatheoretical beliefs and assumptions about the very nature of *proper* scientific values and orientations necessarily inform the evaluation of any given theory. Therefore, yet another critical activity of metatheorizing involves applying these ideas about good science to the analysis of theory and method itself. When scientists evaluate the cogency and coherence of theories and of the hypotheses and conclusions drawn from them, they are drawing upon a set of metatheoretical principles for judging the scientific appropriateness of theory construction and methodological implementation. When scientists ask if a theory is logical, internally consistent, parsimonious, falsifiable, ecologically valid, or consistent with other accepted theories within a scientific domain, they are appealing to a set of metatheoretical presuppositions for what makes good theory. Whether any of these questions make sense in adjudicating theory is an issue that itself derives from metatheoretical presuppositions. For example, some scientific debate exists over the utility and importance of parsimony in theory construction. Such debate reflects metatheoretical division between worldviews (such as the Cartesian-split-mechanistic paradigm) that see all complexity as reducible to basic elements and substances and worldviews (such as the process-relational paradigm) that embrace the importance of contextualizing phenomena within different levels of organization and analysis.

Revealing, examining, or constructing the criteria that scientists and philosophers alike use to formulate and answer questions about the viability of scientific theory and practice constitutes an important form of metatheorizing. Part of these criteria apply not just to single theory building but also to the integration of different theories. Are the elements from multiple theories consistent with each other? What integrated theoretical framework emerges by combining theories? What unique questions might this new framework compel or answer?

Importantly, the practice of building more inclusive theoretical frameworks from existing theories is itself conceptualized differently within the realm of metatheorizing. For some scientists and philosophers, metatheory operates on a qualitatively different level than theory does. Although efforts to integrate different theories, like efforts to evaluate individual theories, draw on a form of metatheorizing, the resulting integrated theories are not themselves metatheories. These higher order syntheses of theories still belong to the domain of theory, tied directly to empirical observation. Scientists have simply used the tools of metatheory to form integrated theories or have reflected on how metatheoretical assumptions and orientations have influenced the process of integration.

For other scientists and philosophers, however, most notably Steven E. Wallis, the overarching theory that results from combining two or more theories itself forms a metatheory. Wallis has argued that one of the key aspects of metatheorizing involves analysis of the common elements of various theories and an integration of those elements into a larger framework. This framework constitutes a metatheory in relation to the theories or elements that it integrates, and the work of metatheory is often centered on the formation of these increasingly comprehensive, discipline-uniting frameworks. Thus, even within metatheoretical circles themselves, debate ensues over where to draw the dividing line between theory and metatheory proper.

David C. Witherington and Timothy I. Vandiver

See also Conceptual Analysis; Cosmology; Metaphysics; Paradigm; Philosophy of Science; Worldview

Further Readings

- Laudan, L. (1977). *Progress and its problems: Toward a theory of scientific growth*. University of California Press.
- Nicholson, D. J., & Dupre, J. (Eds.). (2018). *Everything flows: Towards a processual philosophy of biology*. Oxford University Press.
- Overton, W. F. (2015). Process and relational developmental systems. In W. F. Overton & P. C. Molenaar (Eds.), *Handbook of child psychology and developmental science, Vol. 1: Theory and method* (pp. 9–62). Wiley.
- Pepper, S. C. (1942). *World hypotheses: A study in evidence*. University of California Press.
- Rescher, N. (1996). *Process metaphysics: An introduction to process philosophy*. State University of New York Press.
- Wallis, S. E. (2010). Toward a science of metatheory. *Integral Review*, 6(3), 73–120.

METEOROLOGY

Meteorology. *See* Climate Science

MODELING

Models are of great importance in the natural and social sciences for investigating phenomena and making causal inferences. A wide range of subjects has been studied using models, including biology, communications, economics, education, geology, political science, psychology, and sociology, among others. Such models differ in many ways, but they all aim to represent reality indirectly: The goal is to scrutinize reality through an artificial system that resembles, in some important aspects, the properties of the world that people are attempting to understand. By focusing on some aspects of the target phenomena, models are able to provide manageable representations that can eventually lead to further understanding of those phenomena.

Representing human behavior by modeling is by no means easy, however, and can be a challenge to various disciplines. Researchers are faced

with a number of modeling problems. For instance, much social science research is based on nonexperimental data, such as obtained by survey or quasi-experiment. Moreover, unlike studies in natural sciences, variables to be modeled are unavoidably measured with error, and failing to take measurement error into consideration yields unreliable statistical inferences. In addition, measurements in social sciences may involve categorical and nominal variables, such as gender and ethnicity, besides the quantitative and continuous variables. Finally, with longitudinal data gaining popularity, new modeling methods are needed to take repeated measures into account. This entry provides an overview of how modeling is applied in different disciplines and daily lives and introduces as examples four commonly used models in social sciences, namely categorical models, linear models, nonlinear models, and structural equation models.

When to Use Models

Models are everywhere in our daily lives, and understanding models can help people to become intelligent citizens, think in a clearer way, better understand data, and better strategize and plan. As the British statistician George E. P. Box famously noted, all models are wrong, but some are useful. It makes sense that models can be wrong, given that they are simplifications and abstractions of reality, but how can models be useful for people to work in a better way?

Modeling is nowadays the language not only of the academy but also of a variety of other domains. People from every discipline are using models to enable them to achieve their goals. Financial analysts on *Wall Street* are using models to forecast a corporation's profitability or determine the benefits of a merger. Biologists are using models of the brain to investigate where there are pathways between the neurons. Sociologists use models to identify effects of people's actions and behaviors, and psychologists model to measure people's attitudes. In political science, people use the spatial voting model to predict voting choices based on candidates' and voters' respective political orientations. Linguists also use models to understand the structure of a language under

study. And in daily life, it is not uncommon to receive advice from one acquaintance and the exact opposite advice from another. One possible way to adjudicate between such opposing counsel is by constructing models that can stipulate the conditions under which it would be appropriate to follow each recommended course of action, respectively. Although the actual application of modeling differs from discipline to discipline, in a general sense, modeling is useful for people to think about and to make sense of the world, and perhaps understand how the world could work in a better way.

How to Conduct Modeling

Modeling is a multistep process that can enable people to think more logically. To construct a model, the first step is to decide what factors matter most in determining the outcome one is interested in. Suppose we want to write a model of where students choose to go to college, and there are over 4,000 universities and community colleges in the United States. These institutions vary in terms of the academic resources, reputation, location, costs, and other factors, which is one aspect of the model. The other aspect is that of the students who apply for acceptance at these colleges and universities. The factors that may be relevant in influencing their application decision could be their standardized test score (e.g., SAT), grade point average (GPA), the subject they want to study, and the cost of tuition and other fees. These are determinants that jointly influence students' final decision of which school to attend. For instance, a high school graduate with a 2.5 GPA may be unlikely to apply to Harvard University, implying that GPA may be a variable predicting the outcome. However, in most of the cases, the outcome is not solely determined by a single factor. It is possible that a student with a 4.0 GPA will choose a less prestigious school than Harvard because the former institute provides them with full funding or may chose Stanford over Harvard because of the location. In short, there are multiple relevant factors that should be considered in modeling to understand a phenomenon. Once these factors are identified, we need to think about and establish the

relationships between those factors and make sense of the data.

Why Conduct Modeling?

Social and natural scientists use models to help them understand data better. First, modeling enables people to understand the patterns of data. For instance, the relationship between two variables could be a straight line or a cyclic curve, which enables researchers to get better sense of a large amount of data and make inferences. Therefore, in addition to knowing the patterns, people make use of models to make predictions. We can use data to construct a model and predict a point value from that model. Suppose we want to know how crime rate is related to residents' poverty level and predict the crime rate of a specific community. If the pattern shown in Figure 1 was obtained from the data collected, we know that there is a positive linear relationship between the crime rate and residents' poverty level. Given the

poverty index of a specific community, we are able to make a prediction of the crime rate at that area.

Prediction is not precise, however, and is susceptible to errors. In that case, people are usually interested in not only the point estimate but also the bounds of that estimate to tolerate estimation errors, which is the third way modeling is used. Suppose a chief financial officer of a company wants to know from an economic advisor what the inflation rate will be 10 years from now to make a long-term financial plan. The advisor can hardly answer such question without a sophisticated economic model. Moreover, the model will not only give a point estimate but also provide a range. It is not feasible for a model to produce an exact estimate in the social sciences, such as inflation rate, because too many contingencies are involved. Such complexity and uncertainty make the bounds more valuable, in that people are able to make a comprehensive decision based on the upper and lower bound. Finally, although in many cases people prefer making predictions with models, there are those who use models to estimate

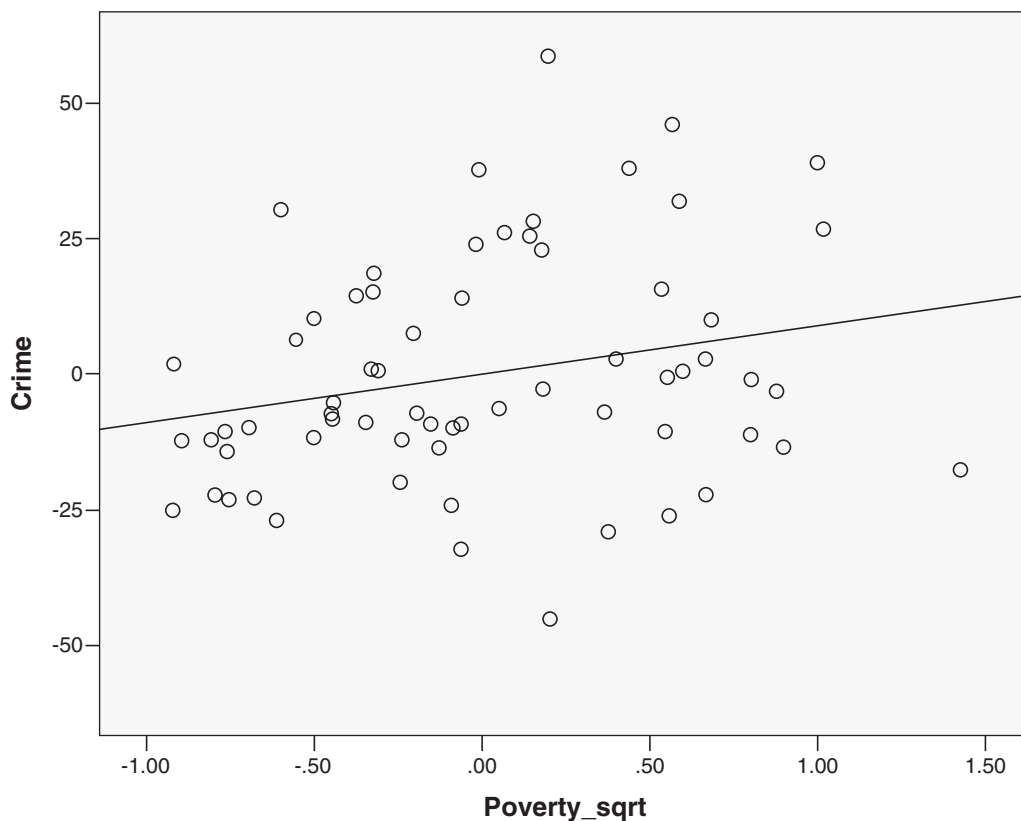


Figure 1 Pattern Between Crime Rate and Residents' Poverty Level

data from the past and to figure out what the past was like. The reason could be that people (e.g., anthropologists, archaeologists, and geologists) sometimes have insufficient access to the data from the past.

In short, modeling enables people to make better sense of data and make predictions for the future and estimates from the past. By doing so, people have better chances to make wiser decisions, strategize better, and design things better. For illustration, let us examine four commonly used models.

Categorical Models

Categorical models, regarded as the simplest of models, are used to classify subjects into different categories on the grounds that such classification can be useful to make better sense of the data. Such usefulness can be reflected by explaining some of the variance in the data set. Here is an example. Suppose that six students took the math exam and got the following scores: Amy scored 100; Ben scored 75; Emily scored 80; Helen scored 90; John scored 65; and Tom scored 70. If we are interested in the variation of these six students' math scores, we need to first calculate the average grade of this group of students, which is $(100 + 75 + 80 + 90 + 65 + 70)/6 = 80$. In calculating the variation, the *sum of squares* (SS) is always applied in statistics to (a) convert negative values into positive ones and (b) to amplify deviations. So the variation is $(100 - 80)^2 + (75 - 80)^2 + (80 - 80)^2 + (90 - 80)^2 + (65 - 80)^2 + (70 - 80)^2 = 850$.

As variation is the part that is left unexplained, a possible way to reduce the variation is to explain why certain subjects have higher values while the other values are lower by categorization. For instance, one way to categorize in this case is by gender. In such case, the average grade and variation for the male group is $(75 + 65 + 70)/3 = 70$ and $(75 - 70)^2 + (65 - 70)^2 + (70 - 70)^2 = 50$, respectively, and for the female group is $(100 + 80 + 90)/3 = 90$ and $(100 - 90)^2 + (80 - 90)^2 + (90 - 90)^2 = 200$. Such categorization by gender substantially reduces the amount of variation left over from 850 to $90 + 200 = 290$. We initially have 850 units of variation that cannot be explained. However, with the categorical model indicating that female students on average

got higher grades than male students in this exam, $(850 - 290)/850 = 65.9\%$ of the total variation was explained. The percentage of variation that is explained by the model is referred to as *R square* (R^2), which falls between 0 and 1. A close-to-zero R^2 indicates that the model fails to explain much variation in the data set, while a larger R^2 shows the models' effectiveness in statistically explaining the data. By knowing that gender explains much variation in students' math grades, further investigation could be conducted to identify the reasons for male students' underperformance and design interventions targeting this group.

In real cases, there are many more subjects than the number in this example, and more categories are needed. For instance, besides gender, a second dimension of categorization can be applied, such as their maternal education, and further categorize them into more groups. Despite different ways of categorization, the primary purpose of categorical models is to explain the variation of data. How to categorize things affects how people think about things, and what decisions they are going to make.

Linear Models

In linear models, there are some explanatory or independent variables typically denoted as X , and outcome or dependent variables referred to as Y . A major activity in linear models is to identify the dependence of Y on X . As noted by Trevor Hastie and Robert Tibshirani, there are three goals of linear modeling: (a) description, to find out the process by which the outcome variables is generated, (b) inference, to detect the contribution of each explanatory variable in predicting the outcome variables, and (c) prediction, to forecast the value of an outcome variable for a specific combination of the values of the independent variables. For instance, public health researchers interested in lung cancer may specify a linear model to describe the process by which lung cancer incidences occur, to infer whether using tobacco increases the risk of lung cancer, and to predict the likelihood to get cancer of a person who uses a high volume of tobacco (X_1), and one who consumes few vegetables (X_2) to get cancer.

A simple linear model can be represented as $Y = \alpha + \beta X$, where Y , the predicted value, is a function of X . α is the *intercept* and β is the *coefficient*. Figure 2 illustrates a typical graph of linear regression that the dependent variable is on the Y -axis while the independent variable is on the X -axis. So the outcome variable can be predicted by $\alpha + \beta X$, and the value of Y equals α when β is zero. β , the slope of this linear model, indicates how fast Y will change if X changes by one unit. Two indicators of β are *sign* and *magnitude*. The sign of the coefficient indicates whether X influences Y in a positive or negative way, and the magnitude represents how much the influence is.

In social sciences, there are seldom outcomes that are predicted by a single variable, but usually, rather, by multiple explanatory variables. In those cases, we need a multiple linear regression model, represented as

$$Y = \alpha + \beta_1 X_1 + \beta_2 X_2$$

Since the regression line or the predicted value can hardly be exactly the same with the value in population but leave part of variance unexplained, which is referred to as *error* (ϵ_i), the model can be represented as follows:

$$Y_i = \alpha + \beta_1 X_{i1} + \beta_2 X_{i2} + \epsilon_i$$

In this model, the coefficients α and β_1 and β_2 are the population regression parameters; the central object of regression analysis is to estimate these coefficients. This model is similar with the Y model in terms of the form as the variables (X_1 , X_2 , Y) and the coefficients we are interested to study are similar; however, the two models are different because in the Y model we focused on what the best line/plane/hyperplane is for a *given data set from a particular sample*. It is rarely true that researchers are only interested in this sample observed. Instead, they want to make inferences of the whole *population*. If we want to predict students' SAT Verbal scores (Y) with their SAT Math scores (X_1) and Teachers' Pay (X_2) in population, the model can be explained like this: The intercept α is the predicted value of Y (SAT Verbal) when both X_1 (SAT Math) and X_2 (Teachers' Pay) are 0. β_1 represents the effect on Y (SAT Verbal) of a one-unit increase in X_1 (SAT Math) holding X_2 (Teachers' Pay) constant. β_2 is the effect on Y (SAT Verbal) of a one-unit increase in X_2 (Teachers' Pay) holding X_1 (SAT Math) constant. As the predicted value Y can hardly be the same as Y_i in reality, ϵ_i represents the aggregated omitted causes of Y in the population, or put differently, the causes of Y beyond the explanatory variable X .

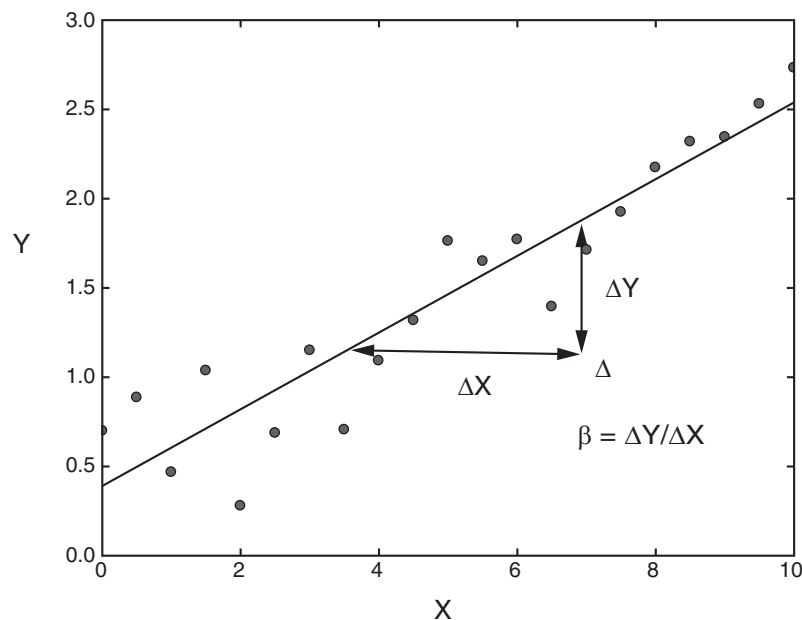


Figure 2 Linear Regression Model

Nonlinear Models

In the real world, many relationships are not linear, and there are many more nonlinear functions than linear ones. For example, the relationship between X and Y can be fitted in a quadratic function, as shown in Figure 3. There are in general two ways to deal with nonlinear models. First, nonlinear models can be approximated with linear functions. Both nonlinear functions in Figure 3 can be approximated by two linear functions according to the slope: The first linear function is from the beginning of the semester to the final exam, and the second function is from the final to the end of the semester. Similarly, regarding cubic, quadratic, and more complicated models, they can also be approximated with a sequence of linear functions. However, analyzing nonlinear models in this way can produce much error and leave considerable variance unexplained. The second way is considered better than the first from this aspect, by including nonlinear terms in the model as follows:

$$Y = \alpha + \beta_1 X + \beta_2 X^2$$

In this model, Y is the function of X and X^2 . As the graph of a quadratic function is a parabola, if $\beta_2 < 0$ (or is a negative number), the parabola

opens downward, and if $a > 0$ (or is a positive number), the parabola opens upward. Instead of having X^2 in the model, the nonlinearity can also be presented in many other ways, such as $\sqrt{X}$, $\ln(X)$, $\sin(X)$, and $\cos(X)$. However, which nonlinear model to choose should be based on both theory (or previous research) and the percentage of variance explained.

For the multivariate nonlinear model, the function can be represented as

$$Y = \alpha + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_1^2 + \beta_5 X_1 X_2$$

α is the intercept representing the average value of Y when $X_1 = X_2 = 0$;

β_1 is the effect of a one-unit increase in X_1 on Y when $X_1 = X_2 = 0$;

β_2 is the effect of a one-unit increase in X_2 on Y when $X_1 = X_2 = 0$;

β_3 is the effect of a one-unit increase in X_1 on the slope relating X_1 and Y ;

β_4 is the effect of a one-unit increase in X_2 on the slope relating X_2 and Y ;

β_5 can be interpreted as either of the following: (a) for a fixed value of X_1 , β_5 is the effect of a one-unit increase in X_2 on the slope of the relationship between X_1 and Y or (b) for a fixed value of X_2 , β_5 is the effect of a one-unit increase in X_1 on slope of the relationship between X_2 and Y .

Structural Equation Models

Structural equation modeling (SEM), also referred to as *simultaneous equation models*, is a comprehensive statistical approach to testing hypotheses about relations among *observed variables*, which are measured, and *latent/unobserved variables*, which are not measured. In an SEM diagram, measured variables are often represented by rectangles or squares, while latent variables are indicated by ellipses or circles. The error terms for measured variables are represented as "E" and the error is referred to as *disturbance* for latent variables, represented as "D." The error terms represent residual variances within variables not accounted for by pathways hypothesized in the model. Structural equation models consist of two parts, a *measurement model* and a *structural model* (Figure 4). The measurement model deals

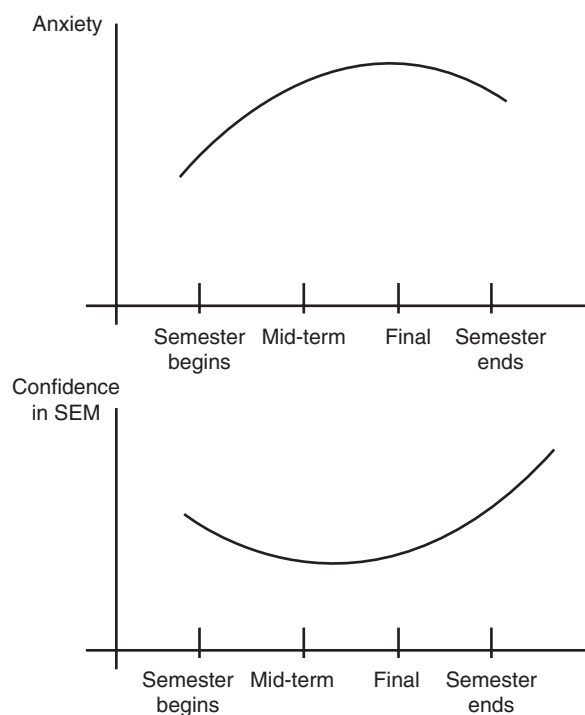


Figure 3 Nonlinear Models

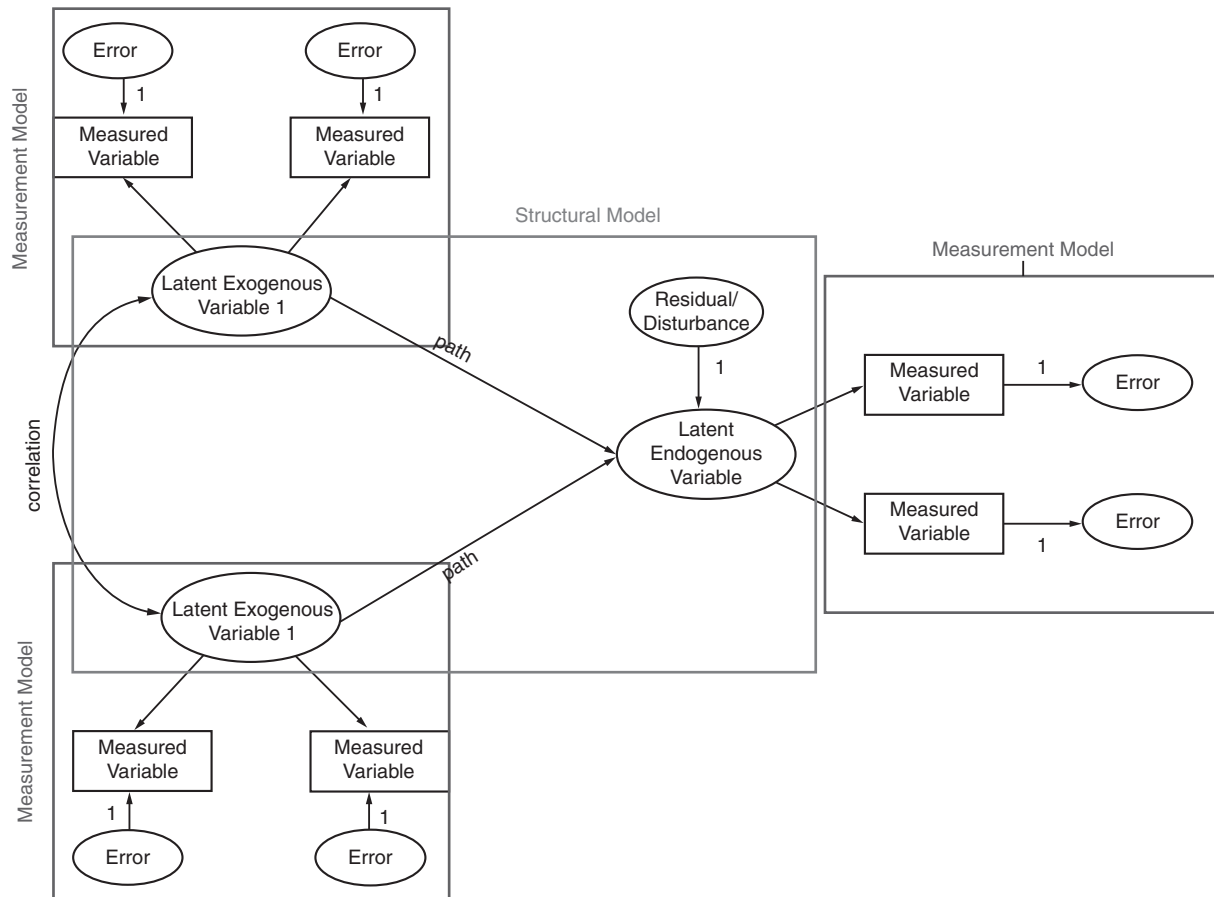


Figure 4 Structural Equation Modeling

with the relationships between measured variables and latent variables, while the structural model only deals with the relationships between latent variables.

SEM is a methodology used to represent, estimate, and test the hypotheses of relationships between variables. Differently from the linear regression models, the variables in SEM are addressed as the *exogenous variables*, which are not caused by any other variables in the model, and the *endogenous variables*, which are explained by exogenous variables. SEM is also used to test hypothesized patterns of directional and nondirectional relationships among a set of observed variables and latent variables. There are generally two goals in SEM: (a) to understand the patterns of correlation/covariance among a set of variables and (b) to explain as much of their variance as possible with the model specified.

There are both similarities and differences between traditional statistical methods, such as the

linear regression models, and the SEM. SEM is similar in several ways to traditional methods like correlation, regression, and analysis of variance. First, both traditional methods and SEM are based on linear statistical models. For instance, the confidants of the *path analysis*, a subset of SEM used to examine causal relationships between two or more variables, can be obtained by running a sequence of multiple regressions. Second, statistical tests associated with both methods are valid only if certain assumptions are met, such as normality; traditional methods assume a normal distribution and SEM assumes multivariate normality. Third, neither linear regression models nor SEM offers to test causality. In other words, no causal relationships can be tested with SEM.

Despite the similarities, SEM has several noteworthy advantages over traditional models. First, unlike traditional methods which specify a default model (e.g., $Y_i = \alpha + \beta_1 X_{i1} + \beta_2 X_{i2} + \epsilon_i$), SEM offers no default model and places few

limitations on what types of relations can be specified. However, the model needs to be estimated and tested by formal specification, which requires researchers to support hypotheses with theory or research and make conceptual sense of the specified model. Second, SEM is a multivariate technique incorporating both observed (measured) and unobserved (latent constructs) variables, while traditional techniques analyze only measured variables. Such technique enables multiple related equations to be solved simultaneously and allows researchers to take into consideration the imperfect nature of measuring observed variables. That the latent variables are assumed to be free of random error is considered as a significant advantage of SEM because error has been estimated and removed, leaving only a common variance. SEM explicitly specifies error, while traditional methods assume measurement occurs without error. Finally, unlike traditional analysis which depends on significance tests alone to determine group differences or relationships between variables, SEM provides no straightforward tests but a number of model fit indices to evaluate the model fit, such as the chi-square test (χ^2), comparative fit index (CFI), Tucker Lewis index, Bentler–Bonett nonnormed fit index, root mean squared error of approximation (RMSEA), and standardized root mean square residual.

There are several rules of thumb to evaluate the adequacy of structural equation models. For instance, the χ^2 test is the test of absolute fit, which indicates the amount of difference between expected and observed covariance matrices. A χ^2 value close to zero with the probability level greater than .05 indicates little difference between the expected and observed covariance matrices. The CFI is the discrepancy function adjusted for sample size. CFI ranges from 0 to 1 with a larger value indicating better model fit. A CFI value of .90 or greater indicates acceptable model fit. RMSEA is related to residual in the model. RMSEA values range from 0 to 1 with a smaller RMSEA value indicating better model fit. Acceptable model fit should be indicated by an RMSEA value of .06 or less, as noted by Li-Tze Hu and Peter Bentler. For more details, see Rex B. Kline (1998).

According to Richard Arminger, Clifford Clogg, and Michael E. Sobel, several circumstances need

to be taken into account on the grounds that modeling in the social sciences is substantially different from that in the natural sciences in several aspects. First, because experimentation is usually not feasible and therefore often replaced by survey, the explanatory variables are not manipulated and well controlled by the researchers, which makes the explanatory variables random rather than fixed. Second, given the complexity of the phenomena in social sciences, there are often quite a large number of possible explanatory variables, compared with the small number of carefully chosen and controlled variables found in the natural sciences. Third, the measurement level of dependent variables is not always continuous but may also be dichotomous (e.g., yes-or-no questions), ordered categorical (e.g., Likert-type scale measuring attitude), or unordered categorical (e.g., choice of preference).

Concluding Remarks

Besides the categorical models, linear and nonlinear models, and structural equation models that have been the focus of this entry, there are many more models used in specific contexts such as multilevel modeling or hierarchical linear modeling, which deals with the nested structure in the data; Bayesian modeling, where a priori knowledge is applied before observing the data; latent growth modeling, which deals with longitudinal data analysis; and others. Although space limitations do not permit discussion of all models in a single entry, readers are encouraged to use the information provided here and the works listed below to explore modeling further.

Qinghua Yang

See also Statistical Inference, Bayesian; Statistical Inference, Frequentist; Statistics

Further Readings

- Arminger, G. Clogg, C. C., & Sobel, M. E. (Eds.). (1995). *Handbook of statistical modeling for the social and behavioral sciences*. Plenum Press.
- Duncan, T. E., Duncan, S. C., & Stoolmiller, M. (1994). Modeling developmental processes using latent growth structural equation methodology. *Applied Psychological Measurement*, 18(4), 343–354.

- Godfrey-Smith, P. (2006). The strategy of model-based science. *Biology and Philosophy*, 21(5), 725–740.
- Hastie, T. J., & Tibshirani, R. J. (1990). *Generalized additive models*. Chapman & Hall.
- Hu, L-T., & Bentler, P. M. (1998). Fit indices in covariance structure modeling: Sensitivity to underparameterized model misspecification. *Psychological Methods*, 3(4), 424–453. <https://psycnet.apa.org/doi/10.1037/1082-989X.3.4.424>
- Kline, R. B. (1998). *Principles and practice of structural equation modeling*. Guilford Press.
- Preacher, K. J. (2008). *Latent growth curve modeling*. Sage Publishing.
- Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods* (2nd ed.). Sage Publishing.
- Singer, J. D., & Willett, J. B. (2003). *Applied longitudinal data analysis: Modeling change and event occurrence*. Oxford University Press.

N

NANOTECHNOLOGY

Nanotechnology is among the most important domains of leading-edge science and technology and is a highly promising field for future research. It deals with both theoretical and experimental structures at the nanolevel (10^{-9} of a meter). The origins of nanotechnology and nanoscience lie in physics, particularly the physics of condensed matter, which is the broadest of the fields of applied physics. Nanotechnology thus falls within the category of the science and engineering of materials. Significantly, in the 1 to 100 nm range, quantum effects come to bear, such as the quantum Hall effect or the Casimir quantum effect, among others.

Indeed, matter behaves differently at the macroscopic level than at the microscopic scale. Thus, a series of discoveries have been made leading to femtochemistry or femtobiology, for instance. The most important levels of the microscopic world exist at nanoscale:

- Mili = 10^{-3} m
- Micro = 10^{-6} m
- Nano = 10^{-9} m
- Pico = 10^{-12} m
- Femto = 10^{-15} m
- Atto = 10^{-18} m
- Zepto = 10^{-21} m
- Yocto = 10^{-24} m
- ...
- The Planck scale = 10^{-33} m

The Planck scale, named for theoretical physicist Max Planck (1858–1947), is the absolute limit in terms of space, beyond which—according to the current state of knowledge—nothing can be assessed. Corresponding to the Planck scale is Planck time which is 10^{-42} s, as the absolute limit in terms of time. The nanoscale stands out as perhaps the most exciting field in which different sciences and disciplines converge, such as physics, chemistry, engineering, mathematics, health sciences, among others.

The counterintuitive behavior of matter at the nanoscale calls attention to the structures of the atom, hence called *nanoparticles*. At very small scale, matter behaves differently allowing us to *see* quantum effects. Those quantum effects are crucial for the understanding of matter in a variety of fields and for working on quantum processes and phenomena.

The term *nanotechnology* was originally coined in 1974 by N. Taniguchi as a process consisting of separating, consolidating, and deforming materials by one atom or one molecule.

Origins

Four steps can be identified regarding the origins of nanotechnology:

- The search for the structure of the microscopic world in the framework of nanotechnology can be said to have begun in 1932, when E. Ruska, a German physicist, and M. Knoll, a German engineer, built the first transmission electron

microscope (TEM). The TEM made it possible to see objects at the molecular level, particularly proteins and the fine structure of metals, for it was possible to see particles smaller than 200 nm by focusing a beam of electrons rather than light through an object.

- On December 29, 1959, Richard Feynman, a leading scientist in quantum physics in general, and in quantum electrodynamics in particular, gave a talk at the California Institute of Technology (Caltech) addressed to the American Physical Society with the title “There’s Plenty of Room at the Bottom,” discussing the theoretical possibility of arranging individual atoms.
- Further on, in the 1970s, Norio Taniguchi, working at the Tokyo University of Science, coined the term *nanotechnology* as the process of manufacturing materials from single atoms or molecules.
- In 1986, K. Eric. Drexler published *Engines of Creation*. The book presents the basis for the most favorable possibilities of research in molecular technology as well as the perils and hopes of the new field. It remains as a landmark in nanotechnology and a source for those in the forefront of the discipline.

Key Developments

Three conceptual, technological, and methodological levels are decisive for the development of nanotechnology: microscopy, synthesis, and assembling. They are closely intertwined. The key term used in nanotechnology at large, however, is *invention*. Indeed, the invention of novel methods, materials, techniques, and approaches comprise the very history of nanotechnology. Briefly discussed below are some of the significant achievements of the people who have played key roles in advancing the path of nanotechnology.

The development of microscopy and its technical refinement has been crucial for the history of nanotechnology. In 1936, the field emission microscope, originally invented by Erwin Müller in Germany, made possible the near-atomic resolution images of materials. In 1981, Gerd Binnig and Heinrich Rohrer invented the scanning tunneling microscope, which allowed for the first time the creation of spatial images of individual atoms. In 1986, Gerd Binnig, Calvin Quate, and Christoph

Gerber invented the atomic force microscope that allows to view, measure, and manipulate materials and forces at the nanoscale.

Based on microscopy, a series of discoveries were made in a range of fields including chemistry, nuclear and quantum physics, electronics, and optics, with surprising and fruitful applications in a variety of domains. Thus, in 1981, A. Ekimov discovered *nanocrystalline*, semiconducting quantum dots in a glass matrix, which opened the gates for further work on electronics and optics with important impact in the health sciences. In 1985, *fullerene* (named after polymath and systems theorist R. Buckminster Fuller), an allotrope of carbon forming a closed or partially closed mesh, in a variety of configurations such as spheres, tubes, and ellipsoids, was discovered by H. Kroto, S. O’Brien, R. Curl, and R. Smalley who worked at Rice University (USA).

To the advances of microscopy, as mentioned, the scanning tunneling microscope and atomic force microscopy continue to deepen, enlarge, and enhance nanoscience research.

During the 1990s, nanotechnology companies begun to be established. To date, the number of companies in the field number more than 100 worldwide. In 1991, S. Iijima discovered the carbon nanotube, a structure that exhibits great properties of strength and electrical and thermal conductivity, fundamental in biological use as well as communications, and photonics. In 1991, C. T. Kresge discovered nanostructured catalytic materials MCM-41 and MCM-48, useful in water treatments, drug delivery, and other applications. Shortly afterward, in 1993, M. Bawendi at MIT invented a method for the controlled synthesis of nanocrystals. Synthesis subsequently became a key method in nanotechnology. Research along this line now influences computing, biology, and high-efficiency photovoltaics.

Along with synthesis, assembling becomes a salient feature of work in nanotechnology. In 1999, W. Ho and H. Lee were able to assemble a molecule with scanning tunneling microscopy. Little by little, similar to what was happening in the realm of genetics, in nanotechnology the work went from just *reading* the atom to *writing* it. Synthesis and assembling become regular methods and approaches that cross disciplines as varied as biology, chemistry, engineering, physics, and environmental studies.

Meaningfully, around the years 2000, several consumer products using nanotechnology began

appearing in the market modifying thus imperceptibly habits and preferences. Tennis rackets, baseball bats, nanosilver antibacterial socks, cell phones, cameras, displays for TVs, therapeutic cosmetics wrinkle-resistant clothing, automobile bumpers, and many other products have been developed until to date that are smoothly but drastically changing everyday life. Nanotechnology is becoming a cultural asset in the life of society. Fifteen industrial sectors can be accounted as leaders in industry and innovation based on nanotechnology.

In 2003, Naomi Halas, Jennifer West, Rebekah Drezek, and Renata Pasqualin developed gold nanoshells for the discovery, diagnosis, and treatment of breast cancer that avoid invasive treatments. The nanoshells are *tuned* so that they can absorb infrared light and identify malign tumors. This is clearly a fundamental path to noninvasive medicine as well as to personalized medicine.

In 2005, Erik Winfree and Paul W. K. Rothemund made possible computations in the process of nanocrystal growth by developing theories of DNA-based computation and algorithmic self-assembly. This offered a way forward to nonconventional computation as well as to biological hypercomputation that aims to shed light on the understanding of life generally.

In 2007, Angela Belcher built a lithium-ion battery with a common type of virus nonharmful to humans. Nanotechnology could thus be identified as one key aspect for the interplay of what has been called the *chip-cell interface*, or the next technological singularity. Such interplay may have important consequences for future research and technology.

The pace of discoveries and achievements is relentless, indeed. Thus, in 2009 and 2010, Ned Seeman helped create several DNA-like nanoscale robotic assembly devices and DNA-assembly lines making programming self-assemble complex processes and structures.

An Expanding Field

Nanotechnology was developed as applied science and it was so conceived for a long time. Subsequently, it has come to shed fresh light on life as well as on the materials under investigation. As such, it has morphed into the larger discipline of *nanoscience*.

Similar to what happened with genetics, the first period of nanotech, from roughly 1959 until the late 1990s, consisted in *reading* subatomic components at the nanoscale. From the 1990s until the present day, the process has included *writing* nanostructures: designing, synthesizing, assembling, and programming various nanoproducts and processes. Furthermore, whereas it was initially a field of interest for physicists and materials engineers, it has since deepened and enlarged its scope, encompassing also biologists, electrical and electronic engineers and computer scientists, environmentalists, and ecologists performing fundamental research on health. Only recently have some social scientists begun to join and nurture the field, thereby further enhancing its cross-disciplinary nature.

When working top-down, the study of nanotechnology moves from mechanical to thermal to optical processes, where sputtering and chemical etching are generally implemented. In other words, nanotechnology engages in the process of modifying materials in order to convert a bulk material into nanoparticles. Correspondingly, the bottom-up procedure uses atomic or molecular condensation that allows a reduction of biological structures or behaviors so that they, in turn, can lead to further chemical or electrochemical reduction. This allows the production of solid material via the sol-gel method and further to vapor deposition and the production of high-quality and high-performance solid materials.

Among the efforts to design, assemble, and program various products are those focused on developing nanorobots, either individually or also as part of swarm robotics. In medicine, they may one day be transported through the bloodstream to specific organs for cleaning up infections, eliminating tumors, or helping the immune system to produce antibodies, for example.

Layers and Materials

Two basic types of nanotechnology materials can be identified: thin films, layers, and surfaces; and multi-layer materials including nanotubes.

Monolayers were the first materials developed and have been implemented chiefly in fields such as chemistry, engineering, and manufacture of a number of devices. Some such materials consist of

one atom or are one molecule deep. Thin films and surfaces are widely produced for coating in general, pharmaceuticals, and in the energy sector.

Regarding multilayer materials, layer graphene stands out for its properties that make possible increased energy capacity of rechargeable batteries as well as enhanced energy storage, flexibility, and lightweight. Current research now anticipates the development and improvement of 2D and 3D structures including silicene (an allotrope of silicon), boron, and molybdenum oxides.

In this second regard, the most conspicuous materials in nanoscience are *nanotubes*, which were the first materials to be developed. Made of carbon, these are cylinders of one or more layers of graphene. Usually, carbon nanotubes range from 0.8 to 2 nm and 5 to 20 nm with lengths from 100 nm to 0.5 meters. Carbon nanotubes are used in biomedical and medical research, energy storage, automotive parts, water filters, coatings, and electromagnetic fields, among other applications.

Quantum dots are man-made semiconducting nanoscale crystals that can transport electrons. When semiconductor particles are small enough, quantum effects emerge, which limit the energies at which the electrons or holes (=absence of electrons) can exist in the particles. Nanocrystals could become an important component for self-assembled nanodevices. So far, the use of quantum dots has been in artificial environments, but some research is being carried out that aims at working also in real-world environments.

Graphene, like fullerene, is an allotrope of carbon consisting of a single layer of atoms arranged in a 2D honeycomb lattice. Graphene was properly isolated and characterized in 2004. It is a zero-gap semiconductor. Among current developments are monolayer and bilayer sheets, graphene superlattices, nanoribbons, graphene quantum dots, graphene oxide, graphene fiber, aerogel, nanocoil, and many others.

Fullerene, as mentioned earlier, consists of carbon atoms connected by single or double bonds that includes a wide family of forms, structures, and shapes named for instance as C_{20} , C_{60} , C_{70} fullerene, and several others.

Nanomaterials are produced in an either of two ways, top-down and bottom-up, always as synthesis processes. In the first case, the production consists of various milling techniques to

crush microparticles. Traditionally, thermal stress is involved and is energy-intensive. In contrast, bottom-up production is based on physicochemical of molecular or atomic self-organization. This method allows for better control of sizes, shapes, and ranges. Hydrothermal, sol-gel, and precipitation processes are regular practices here. Flame reactors working at temperatures around 1,200 to 2,000 °C along with plasma and laser reactors are used in industrial processes for the production of nanomaterials.

Nanorobotics is the process by which several procedures have been developed, such as nanofabrication, which produces nanomotors, nanoactuators, nanosensors, and the modeling of physical materials and processes at nanoscale. Nanorobotics comprises assembling nanometer-sized parts, the manipulation of biological cells or molecules. Rightly understood, nanorobots are nanomotors that are currently used in medicine, environmental studies, and electronics, for instance, to clear bacteria and toxins from blood or waters. Those nanomachines are activated by UV-light, DNA-origin-based nanorobots, light-induced nanotransducers, magnetic nanolink nanoswimmers, and other mechanisms and techniques. Usually, it is harder to produce nanorobots and nanomaterials than acting upon them or making them behave in a variety of ways.

Applications

Among the fields where nanotechnology is being both developed and applied are nanobiotechnology, nanoelectronics, nanoplasmonics (research on optical phenomena at nanoscale), food and agriculture, energy, environmental preservation; various methods and techniques in biomedicine aimed at fighting HIV, cancer, Alzheimer, and food-borne illnesses; the cement industry, construction, displays, furniture, and cosmetics.

Controversy has arisen over human ability and the use of nanotechnology to gain an advantage in sports; the practice is commonly referred to as *technology doping*. Some sports equipment where nanotechnology is being incorporated include tennis and badminton, golf, cycling, skiing, bowling, and archery. The materials used therein encompass carbon nanotubes, silica nanoparticles, fullerenes, carbon nanofibers, nanoclay, and nano-titanium.

In cosmetics, nanotechnology is being incorporated as UV filters and in drug delivery agents. In the automotive industry, this technology is currently present in nanovarnish, carbon black in tires, solar cells, nanofilters, fuel additives, fragrance in the passenger cabin, antiglare coatings, ultrathin layers, and electrochromatic layers, for instance.

As some observers and critics have pointed out, nanotechnology has gradually become more pervasive in everyday life, even if many people do not yet recognize it. The continuing growth of these developments will surely continue making nanotechnology and nanoscience a most salient development in the history of humankind.

The Kavli Prize in Nanotechnology

The Kavli prize awards breakthroughs in science at the atomic scale. It is awarded every 2 years and it was established in 2008. It can be considered as

the equivalent of the Nobel Prize in nanotechnology and nanoscience (Table 1).

Significantly, various awardees of the Kavli prize either had already been awarded, or will be awarded, the Nobel Prize, which is a clear sign of both the work's prestige and scientific rigor.

Science, Economics, and Politics

Finally, as it is well known, there is a scientific and political race around nanotechnology very much as it happened in the past with the exploration of outer space. The two leaders in the race are the United States and China, and the subject around which the race is defined is the course of the fourth industrial revolution in general and the role of artificial intelligence and the Internet generation 5 and 6 (5G and 6G) networks. To be sure, nanotechnology stands at the very center of the current and possible outcomes of artificial intelligence and the fourth industrial revolution. It

Table 1 Winners of the Kavli Prize in Nanoscience, 2008–2020

<i>Name</i>	<i>Country</i>	<i>University or Laboratory</i>	<i>Year</i>
Louis E. Brus	USA	Columbia University	2008
Sumio Iijima	Japan	Meijo University	2008
Donald M. Eigler	USA	IBM Almaden Research Center	2008
Nadrian Seeman	USA	New York University	2010
Mildred S. Dresselhaus	USA	MIT	2012
Thomas W. Ebbesen	France	University of Strasbourg	2014
Stefan W. Hell	Germany	Max Planck Institute	2014
Sir John B. Pendry	UK	Imperial College	2014
Gerd Binnig	Switzerland	Former member of IBM Research Laboratory	2016
Cristoph Gerber	Switzerland	University of Basel	2016
Calvin Quate	USA	Stanford University	2016
Emmanuelle Charpentier	France	Max Planck Institute	2018
Jennifer A. Doudna	USA	University of California	2018
Virginijus Siksnys	Lithuania	Vilnius University	2018
Harald Rose	Germany	University of Ulm	2020
Maximilian Harder	Austria	European Molecular Laboratory in Heidelberg	2020
Knut Urban	Germany	Max Planck Institute	2020
Ondrej Krivanek	UK and Czech Republic	R&D director at Gatan Inc., USA	2020

should be mentioned that the technologies and capacities of and around nanotechnology are extremely expensive, which means that a gap will be rapidly produced between those countries with human, technical, and scientific capacities and those countries that cannot afford the costs of the competition.

Indeed, an atomic force microscope, for instance, can easily cost over half a million dollars. Of course, a good laboratory working on nanotechnology needs more than that.

Currently, there are over 1,200 research groups, laboratories, and departments in universities worldwide. In 2020, the number of private industry and private laboratories exceeded 240. There are over 340 national and international initiatives and networks. Specifically working on raw materials, there are more than 310 groups and laboratories. Currently, there are nearly 200 groups, both private and public, focused on biomedicine and the life sciences. More than 900 companies—baseball bats, cosmetics, and so on—exceeds 900.

Furthermore, regarding education and research on nanotechnology, as of 2020 there are 67 B.A. degree programs, 157 master's degree programs, and 44 doctoral degree programs, which clearly shows both the importance and the trend of education and research in leading-edge science on nanotechnology. The number of journals either particularly focused on nanotechnology or that accept publications in the field has risen to 225, further illustrating that nanotechnology is rapidly becoming a major player in mainstream science.

Carlos Eduardo Maldonado

See also Engineering, Contemporary; Medicine, Contemporary; Physics, Solid-State

Further Readings

- Agrawal, D. C. (2013). *Introduction to nanoscience and nanomaterials*. World Scientific.
- Anis, M., AlTaher, G., Sarhan, W., & Elsemary, M. (2017). *Nanovate. Commercializing disruptive nanotechnologies*. Springer Verlag
- Brushnan, B. (Ed.). (2012). *Encyclopedia of nanotechnology*. Springer Verlag.
- Contera, S. (2019). *Nano comes to life: How nanotechnology is transforming medicine and the future of biology*. Princeton University Press.
- Drexler, K. E. (2013). *Radical abundance: How a revolution in nanotechnology will change civilization*. Public Affairs
- Drexler, K. E. (1986). *Engines of creation: The coming era of nanotechnology*. Anchor.
- Durkan, C. (2019). *Size really does matter: The nanotechnology revolution*. World Scientific.
- Editorial. (2020). In recognition of aberration-corrected TEM. *Nature Nanotechnology*, 15, 417.
- Editorial, (2018). Bacterial nanotechnology. *Nature Nanotechnology*, 13, 435.
- Feynman, R. (1960). There's plenty of room at the bottom. *Engineering and Science*, 23(5). <http://www.nanoparticles.org/pdf/Feynman.pdf>
- Guston, D. H. (Ed.). (2010). *Encyclopedia of nanoscience and society*. Sage Publications.
- Hell, S. W. (2003). Toward fluorescence nanoscopy. *Nature Biotechnology*, 21(11), 1347–1355, <https://doi.org/10-1038/nbt895>
- Huang, R., & Wang, Y. (Eds.) (2020). *New nanomaterials and techniques for tumor-targeted systems*. Springer Verlag.
- Khudyakov, Y. E., & Pumpens, P. (2016). *Viral nanotechnology*. CRC Press.
- Lourtioz, J.-M., Lahmani, M., Dupas-Haeberlin, C., & Hesto, P. (Eds.) (2016). *Nanoscience and nanotechnology. Evolution or revolution?* Springer Verlag
- Postorino, S., Grassano, D., D'Alessandro, M., Pianetti, A., Ranjan, S., Dasgupta, N., & Lichtfouse, E. (Eds.). (2016). *Nanoscience in food and agriculture* (Vol. 1). Springer Verlag
- Rogers, B., Adams, J., & Pennathur, S. (2014). *Nanotechnology: Understanding small systems* (3rd ed.). CRC Press
- Sengupta A., & Sarkar, C. K. (2015). *Introduction to nano: Basics to nanoscience and nanotechnology*. Springer Verlag.
- Urban, K. W. (2008). Studying atomic structures by aberration-corrected transmission electron microscopy. *Science*, 321(5888), 506–510, <https://doi.org/10.1126/science.1152800>
- Weisman, R. B., & Kono, J. (2019). *Handbook of carbon nanomaterials*. (In two volumes. Volume 9: Optical Properties of Carbon Nanotubes, Part I: A Volume Dedicated to Memory of Professor Mildred S. Dresselhaus. Volume 10: Part II: A Volume Dedicated to Memory of Professor Mildred S. Dresselhaus). World Scientific.

NATURAL RESOURCES

A complex picture of how humans interact with natural systems emerges from the study of, and theorizing about, natural resources. All humans depend on our environment to some degree in satisfying our daily needs for food, shelter, and clothing. To gain those basic necessities entails that some residue is generated from processes related to natural resources extraction, exploitation and consumption. Beyond those basic facts, the amounts and forms of natural resource use vary greatly between locations, socioeconomic tiers, ethnicities, regions, cultures, and countries. Natural resources also present a particular challenge to the theoretical and practical toolkit of several overlapping disciplines such as geography, politics, law, economics, ecology, and environmental studies. One reason for this challenge is difficulty in accounting for all the resources used. Broadly defined, natural resources consist of any natural material considered useful or desirable to someone somewhere; but this can and often does change with location, scarcity, or cultural values. Sand, for instance, is a nuisance in Nevada, where the supply seems endless, but is exported to glass factories in Kentucky as a raw ingredient for glass manufacture. Conversely, water in desert regions can make the difference between life and death but has almost no value as a commodity in regions such as the American Southeast, where water is abundant. While we all need to use resources for food, shelter, and clothing, we consume far more resources indirectly through complex pathways between source and user. Thus, natural resources and how to acquire, store, process, exploit, distribute, buy, sell, trade, steal, divert, preserve, and conserve them—as well as the need to grapple with the residue of resource use—have demanded much time, thought and energy, as discussed in the sections that follow.

Types of Resources

Natural resources usually are considered in different ways, based on the broad categories of renewable and nonrenewable resources. The *non-renewable resources* are essentially finite. They

require an extremely long time to regenerate, which far exceeds human life spans. Therefore, in terms of human use, they are finite in their supply. Nonrenewable resources include the fossil fuels such as oil, coal, and natural gas as well as most metals. Because it takes millions of years for nature to create gold, once the amount of gold in a seam has been mined out, it will not regenerate, and the mine is considered empty or depleted. This is not the case for *renewable resources*, which can be regenerated within the life span of humans or, in some cases, they never deplete at all. Examples include surface water, forests, and most food production. For instance when forests are cut down, if trees are replanted then the forest can regenerate in a matter of decades. Similarly, taking water from rivers decreases the immediate quantity of the resource, but snow melt and precipitation will regenerate the fresh water over the seasons of the year. Within the renewable resources category a special subset exists known as *perpetual resources*, which exist whether we use them or not, and do not diminish with human exploitation. Perpetual resources include wind, solar, and geothermal resources. For example, capturing wind using a wind turbine will not decrease the amount of wind that blows; similarly, no matter how many solar collectors capture solar energy, it will never decrease the intensity or amount of incoming solar radiation. Resources are classified as *mobile resources* if they tend to change location or do not conform to notions of political boundaries or ownership. Migrating fisheries or herds of wild animals would fall into this category. This can also describe the negative impacts or residue of resource use such as air pollution or water pollution. Clouds of smog do not stop at customs checkpoints and do not conform to notions of political boundaries or ownership. *Common pool resources* (also called common property resources) are owned by all in theory even if accessed or utilized by only a few—the atmosphere, sunshine, or the oceans. The last classification, *potential resources*, is that category of things not yet considered to be resources but that could potentially become resources in the future. This category, by its very definition, remains elusive of quantification since the potential uses of resources for future generations will depend greatly on choices made

by today's generation, as well as new technologies, applications, and innovations. An example with positive implications would be the new emerging technologies that allow for the capture of methane for use as energy from decomposing trash. Today's landfills will likely become the next generation's sites of reclamation and mining, for the materials we toss away. On the negative side, species of plants and animals that become extinct can clearly never have the potential to become a natural resource for future generations. It should be noted here that these categories, although useful, are not always mutually exclusive. Groundwater, for instance, could be considered a renewable resource as long as the rate of withdrawing water from an aquifer does not exceed nature's ability to recharge that aquifer. Precipitation drops water on the land, some of which soaks into the soil, percolating down until it replenishes aquifers, within a span of months to years as part of the hydrologic cycle. However, we also have massive stores of groundwater that were the residue of melting as the last ice age receded roughly 13,000 years ago. This is essentially a hydrologic gift from a previous era that will not be replenished in our lifetimes. Exploiting that groundwater, as we currently do in semiarid agricultural regions in the Midwest, would be considered exploiting a nonrenewable resource.

Resources and Dominance

Throughout history, acquisition and control of natural resources has been the cornerstone of much economic, political, social, and foreign policy. The need to acquire and control resources has driven technological innovation as far back as the advent of the water wheel and the fence. From the beginning of human history, resources and the need or desire to acquire or dominate them form the basic logic of warfare, political maneuvering, and statecraft. One prime motivation for war has always been to secure one's own resources (defense) or plunder the resources (offense) of others. In a more productive direction, a core *raison d'être* for governments is to provide security to their citizens, often against armed invaders who seek to acquire natural resources. Another central focus of government is that of developing and exploiting the natural resources within its territory, though in modern

times conserving resources has become more of a concern. Throughout history, countries and empires assessed their strength in three basic metrics: military might, economic prowess, and amount of natural resource wealth. Clearly, advantages exist for those who control resources and disadvantages exist for those who do not. This reality has driven worldwide trends in geopolitical dynamics such as early global trade routes like the Silk Road; the rise of colonialism; global commodity chains; and multinational corporations and globalization. The variations in geographic distribution of natural resources around the planet form the stage upon which these tensions are played out and helps explain why there exist fundamental inequities between the developed world and the developing world—patterns that are commonly referred to as the *global north* and *global south*.

Capitalism and Accumulation

The idea of commerce and trade for profit goes back thousands of years. But not until the convergent evolution of capitalism, paper currency, and changes in political theories of value and property did human see the massive transformation in the political economy and concomitant use of resources that now exist in modern capitalism. Feudalism as a system of land tenure, social organization, and economy dominated for more than a thousand years before its dismantling, which began in the 16th century. In feudal regimes, the currency of power rested in the ownership of land that was exclusively owned by the nobility. The advent of mercantilism and the rise of paper currency allowed for increased accumulation of wealth, without the need for large tracts of land or the physical possession of metals such as gold or silver. Then, fundamental shifts in political and economic theory radically transformed how resources were valued, exploited, and consumed. As agrarian capitalism gave way to mercantilism, and then industrial capitalism, theories from philosophers John Locke (1632–1704) and Adam Smith (1723–1790) underpinned the coevolution of political and economic systems.

Smith challenged fundamental beliefs of mercantilism such as the idea that the world's wealth remained constant; therefore, accumulation of one state's wealth could only happen at the expense of

another's loss of wealth. Another theoretical foundation is Locke's theory regarding labor, which states that goods produced by nature have little or no value until labor is added to them. Locke posited that adding labor generates value for goods, and property is created through the application of labor. Locke also argued that said property derived from the application of labor was a natural right. Unlike the economist Smith, Locke did believe there were limits to accumulation. According to Locke, labor creates property, but man's capacity to consume inherently limits man's capacity to produce; any excess would just spoil, so why produce more than can be consumed? With the introduction of durable goods, however, this limit of consumption faded—as did the idea of limits to accumulation. Upon these twin pillars established by Smith and Locke, the fundamental ideas of industrial capitalism and private property were established, and it is this system that has dominated economic thinking for the last 200 frenetic years of expansion. Add to this the ideas of Jeremy Bentham (1748–1832) and John Stuart Mill (1806–1873) and the result is a system predicated on the belief that the greatest good for the greatest number arises when each individual maximizes his utility given the economic options at his disposal. Maximizing utility centers on maximizing profits, and that instinct drives the laws of supply and demand. It is upon these deeply embedded ideas that industrial capitalism feeds the desire to continually accumulate more wealth. Of course, natural resources form the basic raw ingredient in this philosophy of political economy. Nowhere in the basic premises of industrial capitalism is there any discussion of limits on accumulation of wealth, or of the consumption of finite resources.

Many critiques of capitalism exist, but the criticisms related to natural resources come most powerfully from Marxist perspectives. Marxists see capital not as a relation between people and things but as a social relation, and they point out that this relation is predicated on sustained exploitation of human labor by the few who own the means of production. Since the core motivation in capitalism centers on making profit, those who own the means of production are rewarded (with higher profits) for exploitation of labor, the environment, or both. From the Marxist perspective, a

system that is built around the belief that there are no limits to accumulating wealth, and that all trade should be unfettered, is fundamentally flawed. It generates a never-ending spiral of growth whereby natural resources are exploited and products are produced, which require more markets that generate more consumers who buy more things—which in turn generates more residue and more inequality, since only a few control the means of production. There is no steady state or equilibrium in a system built on inequality, predicated on endless exploitation of natural resources. But in fact we live in a finite world with finite natural resources.

Scarcity, Opportunity Cost, and Externalities

Economists find natural resources particularly slippery as an element of economic systems, and their attempts to grapple with natural resources provide us with useful ideas related to *scarcity*, opportunity cost, and externalities. A cornerstone of economics rests on choices and value. When an item is scarce its value will rise; thus as supply dwindles, demand rises and changes in prices reflect this shifting equilibrium. But the inability to determine scarcity or to count the amount of many natural resources skews all attempts at prediction of future trends. *Relative scarcity* occurs either with uneven distribution of resources, or when one group controls and limits access, ownership, or distribution of a resource—usually at the expense of other users. *Absolute scarcity* means depleting a resource to the point of exhausting it completely. As an example, oil is geographically concentrated in just a few places, so it is relatively scarce, and most developed nations import oil as an energy source. In the 1970s, Organization of the Petroleum Exporting Countries imposed an oil embargo restricting access to their oil, creating relative scarcity and pushing oil prices up. But on a planetary scale, oil is also a nonrenewable resource and subject to absolute scarcity as we exploit more and more of it. Knowing when that point of scarcity is reached can be extremely difficult to pinpoint as new reserves might be found (or not), new alternative fuels come into use, and new technologies change the rules of the game.

Economists developed the concept of *opportunity cost* as a metric for quantifying the consequences of choices. Opportunity costs arise when choices must be made among limited options, and choosing one option means an individual loses the opportunity to choose the other. Thus, if an individual chooses to go to university, that entails the lost opportunity of 4 years as an employed worker. Economists compare the relative merits of these two pathways by estimating associated costs and benefits of each path, such as higher income from future employment after university compared with lower income but more earning years. Natural resources elude such calculations because some resources we cannot live without, like water, while others form luxuries. Similarly, in some cases one resource can be substituted for another, but not in all cases. Water is needed to live, and no substitute will suffice if there is no water. Natural gas may be substituted for coal as a source of electricity as market prices vary, but both are nonrenewable resources. When they are gone, what then? Also, economists struggle to quantify lost opportunity costs associated with *not* using a resource. There is no easy way to quantify the opportunity cost of not cutting a forest, or not depleting a shared fishery or not polluting air versus cutting forests, overfishing, or polluting air.

Externalities pose similar problems. These are unaccounted for or unintended impacts from other activities that change hands without agreed upon prices, and often without knowledge or consent. This can be positive or negative, but negative impacts dominate economic debates. Water pollution is an externality that results from many industrial and agricultural activities, but it is not accounted for in the cycle of production, distribution, and consumption of products. Economists can quantify the price of land, seeds, fuel for tractors, storage and distribution of grain, prices of produce, and so on, but the resulting pollution in the waterways from agricultural runoff is not taken into account or quantified. It has no price, but clearly there is a cost. Usually externalities are disproportionately felt, but not by the producer or consumer. When a car factory pollutes a nearby river, it is not the manufacturer of cars who feels the impact, but completely unrelated people downstream. The inability to establish cost or value or price makes natural resources and polluting residues a gap that

resource economists still struggle to address. This imbalance in impacts of natural resource use or externalities resulting from resource exploitation has spawned the new interdisciplinary concept of *environmental justice*.

Malthus, Overpopulation, and Neo-Malthusians

Prior to industrialization, human population tended toward steady but slow growth. With births mostly canceled out by deaths, global population took nearly 50,000 years to reach the mark of 1 billion people, achieved in the early 1800s. However, after industrialization began, population growth rates skyrocketed and continue to do so today—with global population doubling to 2 billion in a century (early 1900s) then again to 4 billion by 1974 and the expectation of doubling again to 8 billion by the end of the current decade. The economist and demographer Thomas Malthus (1766–1834) first highlighted the dangers of overpopulation by connecting population growth rates with far less rapid growth rates in food resources as he studied the overcrowded slums of London. He famously noted as early as 1798 that population was increasing geometrically (or as we would now say, exponentially), while food production can only increase arithmetically. Simply put, population increases too rapidly compared with the rate of man's ability to increase food supplies, so population will outstrip food resources, and war, poverty, and famine would be inevitable at the point where the availability of food could not keep up with the need for it. His early theory on resource limitations remains embedded in our lexicon today; those who apply Malthus's theories to the current state of the world are known as *neo-Malthusians* and they share fundamental agreement that the Earth's agricultural capacity can only sustain a limited population. Although the exact number of Earth's carrying capacity is not known, we clearly will reach or already have reached beyond the population numbers that can be sustained; thus, overpopulation leads to persistent poverty and social degradation.

Malthus lived before the discovery of fossil fuels, the internal combustion engine, and electricity. These innovations have propelled rapid technological innovation that phenomenally increased

our capacity to grow food. Because the global population has massively increased and the dire predictions of Malthus have not yet occurred on a global scale, the debate continues, unresolved. Some suggest that the fulfillment of Malthus's predictions may merely have been delayed but are still correct and inevitable, and calamity is sure to occur at some point. Others believe Malthus's predictions will be avoidable because technological improvements will always keep pace with increased need for food. Neo-Malthusians have always been controversial, leading to much debate and dispute. One major criticism of neo-Malthusians' arguments points out that food distribution, not the production of food, remains a leading factor in persistent poverty. Thus, poverty would exist in some circumstances no matter the total population. Another criticism suggests that people will adapt to their changing environments and find alternatives to the depleted resources, be they food resources or other resources. The debate continues and now has expanded beyond food to include the wider reality of many finite resources, but the question remains as to whether Malthus was correct.

Tragedy of the Commons

In 1968, Garrett Hardin, an ecologist, set off a storm of debate across several fields of study with an article in the journal *Science* entitled "The Tragedy of the Commons," in which he reconsidered the social question of population issues, but with a focus on economic theory and natural resources. His argument continues to be a theoretical touchstone for much of the environmental and natural resource community, while its implications echo outward to many disciplines. According to Hardin, when individuals have unlimited access to natural resources, they will inevitably act in their best interests to maximize their own utility, that is to say, maximize their ability to *utilize* natural resources to satisfy needs or wants. If each individual profits by exploiting as much of the resource as possible, but the group as a whole shares in the negative utility caused by depletion of the resource, then each individual would be compelled to overuse the resource because the individual gains all of the profit but shares only a fraction of the impact felt from the degraded resource.

In his article, Hardin illustrates his point using the example of common grazing lands for herd owners. Each individual will seek to maximize their own profit from the resource, so it would be in each owner's best interest to add one more sheep to their herd. Every herd owner does the same thing, and the logic appears sound. Each individual herd owner increases their own herd, thus their profits, and every individual gains wealth. The dilemma comes when the common resource, in this case the common grazing area, becomes degraded to the point that it can no longer be used at all, and the grazing economy collapses. In fact, the herders are locked into a system that compels each individual to grab as much of the resource as possible at ever-increasing rates. In the case of the herd owners, if an individual does *not* add more sheep to their own herd, then other competing individuals will simply add more, and they will miss the opportunity to maximize utility. The hard-hitting message contained in the example rests with the recognition that it is possible for each individual to gain by acting in their own self-interest, while the group as a whole loses in the end.

This exposes a dangerous but previously unaddressed problem at the heart of capitalism and the neoliberal economy. One of the fundamental tenets of neoliberal economy assumes that all actors seek to maximize their own utility and should behave accordingly, and this is the basis of rational actions. But, if all actors behave rationally, then they will seek to maximize their utility in the market, even as the resources upon which the economy depends become degraded then depleted, and eventually disappear completely. This leads to total economic and ecologic collapse, precipitated by the destruction of the common pool resources. Hardin applied this logic to the threat of overpopulation, viewing the planetary carrying capacity as the commons, and each new person to feed as a new drain on the common resources of the planet, and concluded that there is no technical solution to this dilemma. In his analysis, Hardin separated out those problems that might have a technical solution, or a change in the techniques or science applied to them, from other problems for which there is no technical solution; this latter category requires a change in our morality or human values to find a solution.

Critics of Hardin, most famously Elinor Ostrom and colleagues, point to counterexamples where a tragedy of the commons was avoided and suggest that it is not as prevalent or as inevitable as Hardin would have us believe. She and other critics point to examples in which cooperative behavior results in collective restraint and sustained use of resources, since it is understood that not only individuals but also the collective benefit from this action. Others criticize Hardin for being historically inaccurate by failing to acknowledge the demographic transition or to distinguish between common pool resources and open access resources.

The tragedy of the commons metaphor aptly illustrates that through unlimited access to common resources, unlimited demand in a world of finite resources leads to over-exploitation. Hardin's theory sparked debate in circles of thought as wide and seemingly unconnected as game theory, psychology, cultural anthropology, and strategic model building, as well as its more obvious links to environmental studies, geography, political science, ecology, and economics. The theory stands as an easily graspable conceptual model for a wide array of resource-related and environmental issues based on the dilemma of finite resources such as water, forests, fossil fuels, and arable land, for which we seem to have ever-increasing numbers of users.

Common Pool Resources

Common pool resource theory suggests that the tragedy of the commons is not inevitable, and under specific conditions or scales, natural resources can be successfully and sustainably managed. Common pool resources, by virtue of their characteristics or size, are expensive or impossible to exclude people from obtaining benefits by using them. Resources such as the air we breathe, or the oceans, rivers, forests, pastures, and most coastal fisheries all fit this classification. If people cannot be excluded from using the resource, then management clearly becomes problematic, but in situations where conditions such as clearly defined boundaries, collective agreement on use, effective monitoring of violators, and effective conflict-resolution mechanisms exist and function well, the chance of sustaining the common pool resource goes up considerably. Ostrom developed a set of requirements needed to achieve

adaptive governance that involves sharing of information, dealing with conflict, creating responsibility among users of a resource for monitoring it, and developing a flexible infrastructure that can address errors and adapt to changing conditions. It should be noted that there are examples of successful management of common pool resources, but only at local scales, with specific resources. This does not lessen the success of the management regime. It does suggest that common pool resources will resist any unifying theory of management and may need to be addressed on a case-by-case basis according to the characteristics of the resource and the users. Critics point out that any management regime will likely become less accurate and meaningful at higher scales of governance. But the number and variety of local traditions of successful self-management or small-scale management of common pool resources do show convincingly that users do not inevitably overexploit natural resources in the absence of either privatization or regulations.

Gaia Theory

Developed by the scientist James Lovelock and the microbiologist Lynn Margulis in the 1970s, *Gaia theory* suggests that the biotic elements of Earth interact with inorganic elements of our environment to form a self-regulating interconnected system that contributes to maintaining the conditions for life on the planet. Named after the Greek goddess Gaia, this concept, also known as the *Gaia principle* and *Gaia hypothesis*, suggests that organisms coevolve with their environment. In other words, biota influences aspects of their inorganic environment, such as temperature or atmospheric conditions, and the altered environment in turn influences the development of biota, and together, these forces generate a stable system described as a *planetary homeostasis*. The interactions are complex, but here are two simplified examples used in Gaia theory: air temperature and atmosphere. Since the beginning of life on planet Earth, there has been a 25% increase in the amount of heat from the sun, but the surface temperature has remained virtually constant. Also, our atmosphere consists of roughly 79% nitrogen, 20.8% oxygen, and 0.03% carbon dioxide (CO₂). This highly unstable mix of reactive gases could

not be maintained without constant replacement and/or removal by the Earth's biota. Earlier in Earth's history, the atmosphere contained higher levels of CO₂, and these CO₂ levels produced a greenhouse effect that warmed the planet. Later, biota on the planet has produced increasing amounts of oxygen and methane, via cycling through oceans and rocks; these changes in the atmospheric chemistry have been maintained by evolving species. Now CO₂ is minimized by pumping down of carbon through the biota. Without the presence of greenhouse gases in the atmosphere, scientists estimate our planet's surface temperature would be approximately 19°C.

The idea of a planetary homeostasis, influenced by biotic factors, already exists and is studied by scholars from academic disciplines including ecology, biogeochemistry, and Earth system sciences. However, a unique element of Gaia theory puts forth the assessment that this homeostatic balance is pursued actively, if unconsciously, to achieve optimum conditions for life on the planet—even if external forces threaten that stability. Thus, taken to its logical conclusion, the theory implies that the biotic balance with the regulatory systems of the planet, such as oxygen in the atmosphere or salinity in the ocean, will maintain balances needed to optimize a biogeochemical balance, even if external human activity threatens that equilibrium. Gaia theorists believe that humans could find themselves one day headed toward extinction as a species, but that the planet as a whole would still remain, with its regulatory systems intact. Gaia theory essentially posits that the Earth is in some complex sense *alive* and functions as an intricately interconnected singular entity.

This theory has generated much speculation and debate across the academic spectrum, from philosophers to Darwinists. Philosophers use Gaia theory to begin asking fundamental questions about teleology, the meaning of *living*, and the need for consciousness in order to act in unison. To achieve this, reconsiderations of basic ethical values or concern for nonhuman and even nonliving components of the natural world would be required. The Gaian worldview rests on the notion of large-scale cooperation rather than competition as a driving force and simultaneously suggests that the planet's natural systems are not only interconnected but also function toward a common purpose—stabilization.

On the other end of the spectrum, staunch Darwinists critique the Gaia theory and suggest that evolutionary strategies could never account for complex interaction among nonsentient biota at such large scales. For Darwinists, there is simply no plausible way that the feedback loops Lovelock says work to stabilize the planet could have evolved. For Darwinists, some mechanism must be shown to control the processes described, and if they see no mechanism at work, they see no larger cohesive end. Gaia theory has also set off arguments among scientists about what it means to be *living*, and Darwinist critics note that since Gaia cannot reproduce itself, it fails to meet one of the key components in the accepted definition of *living* among the scientific community.

Resource Law

Resources also pose complications for the legal system. In the global East, resources typically come bundled with property rights, while in the West states, resources and property are severed. This means that property owners in the East also simultaneously own mineral rights, water rights, and forest rights on that property, while in the West, it is common for property to be owned by one person but, say, mineral rights owned by another. This is especially true with water rights traditions. In the East, *riparian law* dominates, which means that water cannot be taken out of its watershed, and only property that touches fresh water has rights to that water. In the West, *prior appropriation* is the rule, which allows water to be moved out of its watershed, and for water rights going to the first to claim them. For this reason, prior appropriation is often known as “first in time, first in right.” These divergent traditions come out of the radically different climate regimes and availability of water in each geographic region.

Gina Bloodworth

See also Environmental Studies

Further Readings

Cutter, S., & Renwick, W. (2003). *Exploitation, conservation, preservation: A geographic perspective on natural resource use*. Wiley.

- Hardin, G. (1968). Tragedy of the commons. *Science*, 162(3859), 1243–1248. <http://www.jstor.org/stable/1724745?origin=JSTOR-pdf>
- Lovelock, J. (1995). *Gaia: A new look at life on Earth*. Oxford University Press.
- Malthus, T. R. (1826). *An essay on the principle of population* (6th ed.). John Murray.
- Ostrom, E., Burger, J., Field, C. B., Norgaard, R. B., & Policansky, D. (1999). Revisiting the commons: Local lessons, global challenges. *Science*, 284(5412), 287–282. <https://doi.org/10.1126/science.284.5412.278>

NATURAL SELECTION

Charles Darwin (1809–1882) and Alfred Russel Wallace (1823–1913) are the codiscoverers of natural selection; between them, Darwin is the chief theorist of the notion. For Darwin, natural selection is a principle, the principle by which each slight variation in organismal traits, if useful, is preserved. Darwin's understanding of selection as a principle is tightly bound to his methodology, one derived from the writings of the polymath John Herschel. Contemporary theorists have sought to differently determine the exact contours of natural selection, treating it not as a principle but as a kind of evolution (in some instances) and as a cause of evolution (in others). How natural selection operated among humans was of immediate concern to converts to Darwinism. In nations given to conquest, oppression and war, ones ruled by individuals given to elitism, racism, and ableism, natural selection provided an explanation, if not a justification, of the extermination of native communities in colonial contexts. It was also the source of a threat, as many elites were convinced that, back home in “civilized” countries, people with worse traits had more children. This conviction led to suspicion of humanitarian endeavors, state interventions to control the reproduction of “defectives,” and ultimately to Nazi atrocities. We stand today capable of editing genes by means of the CRISPR gene editing technology, giving us the potential to control human evolution with a precision unimaginable to the breeders of Darwin's time.

The Principle of Natural Selection

Darwin's work, *On the Origin of Species by Means of Natural Selection, or the Preservation of Favoured Races in the Struggle for Life*, aims to explain that “mystery of mysteries,” how species come into existence. Geologists of Darwin's day knew that species came into existence and became extinct at specific moments in the history of the Earth. What makes a species appear, then, is a puzzle. Respectable thinkers in Darwin's day defended creationism, the view that species come into being as a result of God's special acts of creation. Darwin aimed to defeat creationism. Species, argued Darwin, come into being through evolution by natural selection. According to Darwin, natural selection also explains adaptations, the adjustments of fit between organisms and their environment.

In the *Origin*, Darwin develops his theory in accordance with the methodology of John Herschel (1792–1871), the preeminent scientist of Darwin's day, and the one responsible for the “mystery of mysteries” remark. Darwin begins the work with a discussion of breeding in order to establish selection as a *vera causa*, that is, a genuine or “true” cause, and the right sort of thing with which to explain. *Verae causae* stand in contrast to a mere hypothesis, what Herschel would have derided as a “mere figment[s] of mind.” Having established selection as a *vera causa*, Darwin then generalizes from the specific case of artificial selection across a variety of other processes involving diminished human control to finally include selection among wild animals, that is, “natural selection,” the critical concept in Darwin's explanation of speciation and adaptation.

In his first chapter, Darwin offers an extended discussion of variations among breeds of domesticated organisms; he pays particular attention to the different races of pigeons, emphasizing the extent of the differences among them. No one thing is responsible for the variation, but Darwin attributes the bulk of the explanatory burden to human beings' power of accumulative selection: Nature gives successive variations; humans add them up in certain directions useful to ourselves. Genuine causes, *verae causae*, are ones of which

we have experience; artificial selection is a genuine cause because people are known to have used it, in domestic contexts, to produce variations as great as those that distinguish species.

Having established artificial selection as a genuine cause of differences among breeds of domesticated organisms, Darwin moves to generalize from this home case, in a stepwise fashion. Other selection practices that Darwin considers include “roguing,” the removal from seedbeds of deviant forms, and “unconscious selection,” by which breeders improve their stock by allowing only their best animals to breed, something even “savages” do. Even bans on the import of foreign animals create selection processes. Darwin ultimately extends the notion of selection to cover processes occurring among organisms in nature.

Organisms in nature vary just as domesticated organisms do; moreover, they must compete with one another in the struggle for existence (and, among sexual organisms, in the struggle for mating opportunities). More offspring are produced than can survive to reproductive maturity; those that reproduce are a selection from the pool of candidate reproducers. The struggle to join the pool results from checks on population increase, in the form of resource limitations, predation, disease, harsh climate, and innumerable other causes of differential reproduction on the part of population members. Many of these causes affect who makes it into the pool of reproducers by simply killing their rivals off. Any heritable variation conducive to success in this struggle will tend to be preserved, “selected,” leading to the accumulation of profitable variations across generations and ultimately to the evolution both of new species and adaptations.

In the *Origin*, Darwin offers imaginary illustrations of this process, the most famous of which involves a population of wolves that evolve adaptations for increased speed that allow them to better catch their fleet prey. Scenarios such as this one aim to exhibit how the theory works; Darwin’s wolf scenario shows how to use the theory to explain. Since Darwin’s time, there have been a great many applications of his theory to real-world systems, from cichlid fish to water snakes, to side-blotched lizards. Naturalists make genuine

explanatory deployments of evolutionary theory. But Darwin’s illustration is not meant to be an actual explanation of how wolves evolved. As noted earlier in this entry, Darwin aimed to explain the origin of species, but Darwin did not make a catalog of explained bouts of evolution, as though to explain all of them, one by one. The world is vast; its history, long. The diversity of life is great and its historical record is sketchy. Explaining the origin of species does not mean actually explaining every instance of evolution ever, at least not in the rigorous manner that researchers have done for cichlids, snakes, lizards, and countless other cases.

Darwin showed how speciation and adaptation *could be explained*, without appeal to supernatural causes. The case for creationism rests in part on the contention that designed-appearing features of the universe, adaptations and speciation events in geologic record, cannot be explained without appeal to supernatural causes. William Paley, whose reasoning Darwin much admired, inferred the existence of a designer on the basis of the complexity and functionality of natural forms. The theory of evolution by natural selection supplies us, to this day, with a framework that makes speciation and adaptation explicable, both scientifically and naturalistically. Darwin showed us how we may rationally resist the creationist argument that the origin of species and adaptations cannot be explained without a designer; rather, even the most seemingly complex and functional of organismal traits are explicable in terms of evolution by natural selection.

Selection as a Kind of Evolution

For Darwin, natural selection is a notion he is dedicated to securing as a genuine cause. He aims to do this by means of a generalization, one starting from artificial selection. Rather than emphasizing the distinguishing features of natural selection, Darwin is especially concerned to connect it with artificial selection under the broader rubric of selection *tout court*. More contemporary theorists have been much concerned, however, with the distinguishing features of natural selection.

In Darwin's wake, theorists have been concerned to specify the domain of application of his theory. Different writers have contended that evolutionary theory applies more broadly than Darwin initially envisaged, prompting the question of how to precisely state the scope of its application.

In a couple of separate publications, Richard Lewontin (1970, 1978) offers two differing statements of the requirements for evolution by natural selection; he aims to specify on what principles rests Darwin's theory. The first set of principles requires the heritability of fitness (Principle 3) and applies to populations:

1. Different individuals in a population have different morphologies, physiologies, and behaviors (phenotypic variation).
2. Different phenotypes have different rates of survival and reproduction in different environments (differential fitness).
3. There is a correlation between parents and offspring in the contribution of each to future generations (fitness is heritable).

The second set of principles requires the heritability of phenotypic variations rather than fitness and applies to species.

It has proven difficult for theorists to adequately state the requirements for selection. Some issues remain indeterminate. Lewontin was probably better off with requirements for evolution by natural selection that apply to populations; members of any given species are sometimes too spread out to be undergoing a single selection process. But if evolutionary theory is to apply to populations, how these are to be circumscribed remains an outstanding issue.

Statements of the requirements for selection also must face borderline cases and seeming counterexamples. Populations in which fitness is not heritable, but in which variants have different mortality rates, will evolve from an even distribution of fitter and less fit types to a stable polymorphism in which the fitter types are more frequent. This is simply because their death rate is lower. This sort of evolutionary change at least looks a lot like evolution by natural selection, though it doesn't meet Lewontin's conditions. Similarly, especially heritable variations will spread in a

population despite not being especially conducive to reproduction: Here, we have evolution by natural selection without differential fitness.

Lewontin wrote that the theory of evolution by natural selection rested on his principles. But he effectively makes evolution a feature of the systems to which the theory applies. Curiously, this may be too strong a constraint. While we might initially think to say that evolutionary theory is about evolution, that position is likely unsustainable.

The contention that evolutionary theory applies to systems that evolve is odd from a historical perspective; Darwin himself used the theory he himself developed to explain at least one polymorphism, heterostyly. Plants exhibiting heterostyly develop two, or sometimes three, different forms of flower whose reproductive organs vary in a number of ways, principally length. Darwin regarded heterostyly as a polymorphism that existed to stimulate the intercrossing of different plants. Darwin theory, at least, may explain cases of nonevolution as well as cases of evolution. Furthermore, the same population genetics equations that can be used to show how diverse sorts of populations evolve may equally be used to show how populations reach stable polymorphisms; only the values of some variables need to be adjusted. Evolutionary theory, or at very least population genetics, is widely applied to natural populations to explain polymorphisms; the stable polymorphism resulting from heterozygote superiority, or overdominance, is merely the most well-known example.

It is not just evolution together with stable polymorphisms that evolutionary theory explains. The side-blotched lizards mentioned earlier exhibit a kind of cyclical dynamics. The lizards come in three different colors and selection works against each type as it becomes more common. Ultimately, the system dynamics explicable using evolutionary theory are diverse: to cycles and stable polymorphisms, we should add protected polymorphisms and even random evolutionary change.

To see this last point, consider how, in less idealized, more "realistic" population genetics models, evolution is treated as a random process, as the result of both selection and drift. Presumably, drift is an element of evolutionary theory alongside selection; both are co-contributors to the

dynamics are any actual populations in nature. But if that is so, then just as selection-only models are a core element of evolutionary theory, drift-only models must equally be such. Indeed, much genetic polymorphism may be due to drift and mutation alone (Motoo Kimura's *neutral theory of molecular evolution*) or almost due to drift and mutation alone (Tomoko Ohta's *nearly neutral theory of molecular evolution*). Indeed, Lewontin himself, along with Stephen Jay Gould, critiqued the adaptationist program dedicated to explaining all traits as adaptations that evolved from natural selection; other mechanisms, including drift along with developmental and phyletic constraints, play important explanatory roles.

Conditions for evolution by natural selection, then, whatever their function, will not suffice to state the conditions of application of a contemporary scientific theory, since the theory that covers evolution by natural selection explains more than just evolution and does so in terms of more than just selection. Other theorists, equally concerned with the scope of evolutionary theory, have approached stating the conditions of its application by means of the notion of replication. Richard Dawkins's *The Selfish Gene* will forever be remembered for pioneering the term *meme*, that is, a unit of cultural information that spreads by imitation, much as genes spread by copying. Dawkins, together with David Hull, singles out active germ-line replicators as crucial to triggering the application of evolutionary theory. Active replicators influence their probability of being copied while germ line ones have potentially infinite lives in the form of copies; nonneutral genes found in gametes are the paradigm, but not sole, cases of such replicators.

Critics of replicator selectionists have countered that copying is not, in fact, necessary for evolution by natural selection; similarity between parents and offspring is sufficient. (Darwin, for his part, knew nothing of replicators or genes and grew to emphasize the importance of continuous variation in later editions of the *Origin*.) Equally, it is arguable that a gene need not be a germ-line replicator in order to undergo selection; in the immune system, a selection process determines cell frequencies. Indeed, theorists have invoked selection processes to explain phenomena even further afield, including quantum mechanics.

Selection as a Cause of Evolution

Recall that, at the end of the discussion of Lewontin's conditions, we considered how cases of random evolution, ones *without selection*, belong within the purview of evolutionary theory. Indeed, this was a point made using the notion of selection and contrasting it with *drift*. So there is at least one other ordinary usage of the notion of natural selection, outside Lewontin's conditions, that applies to a component of the evolutionary process: Selection contributes to determining the dynamics of natural systems but does so along with other phenomena, including mutation, migration, segregation, and drift, where each of these is quantified separately in formal models of system dynamics. On this approach, selection is quantified by fitness coefficients, while other phenomena (how inheritance works, which types of descendants are produced by which types of ancestor, etc.) are captured other components of formal models. Selection here is *distinguished* from inheritance and reproduction, rather than being brought into being (in part) by these, as in Lewontin's conditions (recall that Lewontin included among his principles of evolution by natural selection heritability and descendant production). When "natural selection" is used to mean *that which is quantified by fitness coefficients*, selection is being treated as a cause of type frequencies, rather than as a kind of evolution, as it is for Lewontin. Indeed, while a system governed by equations featuring fitness variables may be said to undergo selection for this reason, this by no means guarantees evolutionary change, even when the fitness variables take different values. A multitude of systems of equations with variant fitness variables determine stable polymorphisms.

Historically, population genetics models of selection have undergone a history of complexification. Early models were simple and, importantly, analytically tractable, beginning with the simple case of two alleles at a single locus. Theorists have since developed ever more complex models suitable for ever more diverse systems, often labeling the newly considered systems and associated models as ones undergoing types of selection. So, we get gametic selection models (for segregation distortion), multilocus selection

models (of variable number of loci), frequency-dependent models (with any variety of function determining fitness values), and variable selection models for different sorts of population structures, to name just a few.

Consider, in particular, spatially variable selection models. These are suitable for systems in which at least one nondegenerately distributed environmental cause interacts with the variant types distinguished in the population. (Two causes interact when the influence of at least one is a function of the value of the other.) The conditions that trigger the application of variable selection models can be stated rigorously, as we did just now, using the same language that is used in quantitative causal modeling elsewhere in the sciences. This restatement condition extends to the rest of the different types of selection. Fitness coefficients are, in the end, coefficients. They quantify the extent of a dependency. In evolutionary theory, they quantify the dependency between being an entity of a particular type and contributions to subsequent generations. Sometimes, this dependency is environment specific in a variable environment, triggering the deployment of systems of equations for population members distinguished by environment *and* genotype (spatially variable selection models); sometimes, this dependency is specific to what other members co-occupy the population (frequency-dependent selection models); sometimes, this last dependency is specific to subgroup (as in group-selection models); and so on. These are different types of selection in name only; the deployment of the associated models is triggered by interactive causes in the causal context of the population.

Darwinism in Historical Context

Darwin's *Origin* was widely read. He drew readers' attention to the potential for evolutionary change in populations, driven by differential rates of reproduction. Even though Darwin would not himself publish *The Descent of Man* until 1871, readers of the *Origin* could hardly fail to wonder which human beings reproduced the most and what evolutionary consequences could be expected as a result.

Thinkers in Darwin's wake were concerned with natural selection among human beings in

two different contexts. The first was the colonial one. Darwin and his contemporaries regarded native peoples as uncivilized, even animalistic, in comparison to Europeans. Europeans were liable to see interactions between themselves and native populations as one-sided competitions between populations. Europeans thought themselves destined to win these contests at the expense of the "savage" races by virtue of their superior moral, intellectual, and physical qualities. The extirpation of native communities was seen as an inevitable, if for some regrettable, consequence of natural selection in the colonial context.

Though a history of evolution by natural selection accounted for the prevalence of socially valuable traits among Europeans (at least as far as they themselves were concerned), Victorians saw themselves as threatened with an evolutionary reversal within their own societies, one that resulted from the greater reproduction rates among the poor. In his 1865 essay, "Hereditary Talent and Character," Francis Galton, Darwin's cousin, argued that human traits relevant to success in society, that is, cognitive, moral, and personality traits, were transmitted from human parents to their offspring. Galton and his contemporaries were convinced, on scant evidence, that the poor and other undesirables—the social residuum—reproduced more than did the higher classes, whose members were, on the whole, more valuable to society. The purported preponderance of unfavorable traits—diminished intelligence, moral character, and industriousness—among those who reproduced the most sparked a widespread worry about decline, one driven by the elitist, ethnocentric, and racist attitudes commonplace in the era. Concerns were especially pronounced given the competition in Europe among rival imperialist nations; a great many British men recruited to serve in the Boer War were rejected as being unfit for military service, something that was taken as evidence that Britain was in decline.

This decline was exacerbated by modern medicine along with humanitarian measures that prevented the elimination from society of the weakest in physical, mental, and moral capacities. Indeed, the moral sentiment to help the desperate presented something of a special difficulty, since the prevalence of such sentiments marked the superiority of European populations over the beastly

savages and yet also threatened that superiority. Even the tendency, on the part of the prudent, to delay marriage and child-rearing until such a time as children could be properly cared for had regrettable, “dysgenic” consequences for the population. Still, Darwin emphasized how, nevertheless, selection remained in some ways a positive force within industrial nations: It improved the body and intellectual capacities, even among the poorest classes. Moreover, those with the worst dispositions were unlikely to foster many children: criminals and the insane got locked up; the violent tended to die violent deaths themselves.

Darwin never advocated for state control over reproduction but was equally not optimistic that the reckless could be persuaded to have fewer children and the gifted, more. By the turn of the 20th century, attitudes had hardened. “Defectives” were segregated to prevent their having children; unconsented sterilization of the same people, a cheaper method of control, was adopted in many European countries and parts of the United States and Canada. No country was more intent on controlling reproduction than Germany; indeed, sterilization programs existed in Germany alongside ever more brutal methods of achieving “race hygiene”; the practices of extermination of criminals and defectives paved the way for the horrors of the Nazi regime.

Eugenics will forever be associated with Nazi atrocities. But our contemporary society lies on the cusp of instituting some very strong selection pressures again human genes, although not by means of coercive, state-level control over which humans have how many children. Direct genetic manipulation of the human germ line is no longer the purview of science fiction, thanks to the recently developed CRISPR technology. Good riddance, say some, to such genes as the one that causes Huntington’s disease; others see a dark future of unequally distributed genetic enhancements. Medicine, once credited with having stalled selection in human societies may soon become a forum for its most exacting application yet.

Laissez-faire opponents of government intervention could find support for their views in Darwin; Darwin himself wrote in the *Descent of Man* that no one could doubt that humanitarian practices were highly injurious to the race of man. But, perhaps surprisingly, thinkers of many political persuasions found support for their positions in

Darwin. Religious skeptics, often political radicals rather than conservatives, appealed to the *Origin* to bolster their opposition to religion. Opposition to war, in which especially a great many fit young men died, could equally be opposed for its deleterious evolutionary consequences. Darwin thought subjugation women’s natural station, but feminists could appeal to Darwinian principles in defense of their views, too. Diminished power and standing for women in society reduced the power of sexual selection to improve the human race: Too many women were forced by their social circumstances to settle for, and have children with, inferior men.

Efforts to make rigorous the notion of natural selection may have outlived their usefulness. The concept is ultimately traced to a theory developed by Darwin to demystify adaptation and speciation, but its initial narrow focus has fallen away. Speciation and adaptation were so important to Darwin because they were mysterious. But what Victorian-era Europeans found mysterious, and what results from a distinct type of process in nature for which there is a definite, particular special scientific theory, are obviously different things. What selection could explain was always going to be more than its initial explanatory uses. So the theory that Darwin used to explain adaptation and speciation could perhaps never have been expected to stay put to explain only within a given purview and only with particular tools. The history of the development of evolutionary theory is one of complexification and expansion. Adjacent to every type of population, ideal or real, that has been modeled by some or another evolutionary theorist is another natural system, just a little bit different, that may itself be modeled by a suitably modified version of the model used for its neighbor. Even features of the environment are now said to persist by means of the process of niche construction, in which organisms modify how natural selection operates upon their populations. A theory that grows by means of iterative extensions and complexifications is not naturally confined to a specific domain; its vocabulary is not a natural target for lapidary regimentation. Given the peculiar history of Darwin’s theory, what exactly counts as natural selection may be neither be determined nor determinable.

Peter Gildenhuys

See also Biology, Evolutionary; Environmental Studies; Genetic Drift; Punctuated Equilibrium

Further Readings

- Baylis, F. (2019). *Altered inheritance: CRISPR and the ethics of human genome editing*. Harvard University Press.
- Bryan, C. (2020). Revising, correcting, and transferring genes. *The American Journal of Bioethics*, 20(8), 7–18.
- Darwin, C. (1988). *On the origin of species by means of natural selection, or the preservation of favoured races in the struggle for life*. New York University Press. (Original work published 1859)
- Gildenhuys, P. (2004). Darwin, Herschel, and the role of analogy in Darwin's origin. *Studies in the History and Philosophy of Science Part C: History and Philosophy of Biological and Biomedical Science*, 35(4), 593–611.
- Godfrey-Smith, P. (2007). Conditions for evolution by natural selection. *The Journal of Philosophy* 104(10), 489–516.
- Gould, S. J., & Lewontin, R. C. (1979). The spandrels of San Marco and the Panglossian paradigm. *Proceedings of the Royal Society of London*, 205(1161), 581–598.
- Gynell, C., Douglas, T., & Savalescu, J. (2015). The ethics of germline gene editing. *Journal of Applied Philosophy*, 34(4), 498–513.
- Lennox, J. G. (1991). Darwinian thought experiments: A function for just-so stories. In G. Massey & T. Horowitz (Eds.), *Thought experiments in science and philosophy*. Rowman and Littlefield.
- Lewontin, R. C. (1970). The units of selection. *Annual Review of Ecology and Systematics*, 1, 1–18.
- Lewontin, R. C. (1978). Adaptation. *Scientific American*, 3, 1–18.
- Paul, D. B. (2009). Darwin, social Darwinism and eugenics. In J. Hodge & G. Radick, (Eds.), *The Cambridge companion to Darwin* (2nd ed). Cambridge University Press.

and neuropharmacology. In the past, neuroscience was mainly associated with biology; nowadays, it is a dramatically growing multidisciplinary area that involves fields as diverse as linguistics, computer science, medicine, philosophy, engineering, physics, and psychology, among others.

Background

Although the way neuroscience is perceived nowadays is connected mainly with the scientific achievements of the 20th and 21st centuries, the origins of neuroscience date back to ancient and even prehistoric times, with the first known trepanation (perforation of the skull) having been conducted as early as the Neolithic era. Apart from its medical dimension, the brain was also the topic of consideration among ancient philosophers. For example, those interested in the philosophy of mind (e.g., Plato) studied the relation between body and mind. During the 18th and 19th centuries, phrenology, which focused on the purported relation between the measurements of the skull and one's character, was popular. Although recognized today as a pseudoscience, phrenology represents the growth of interest in the brain and its relation with other functions of the human body. The most visible progress in neuroscience took place in the 20th and 21st centuries, characterized by significant technological and technical advancements. As in the other fields of science, the development in technology has led to more sophisticated methods of research. Nowadays, neuroscience relies on—for example—magnetic resonance, ultrasonography, and radiology in studying the ways in which the brain and the nervous system function. Moreover, by applying invasive or noninvasive diagnostic methods, neuroscience provides information on the neurologic performance of healthy and unhealthy subjects. Patients whose disease or injury makes invasive methods impractical can often be diagnosed and treated by noninvasive methods. In addition, the rapid and sustained growth of the internet has contributed to knowledge awareness and innovation diffusion of neuroscientific achievements, with online scientific databases, social networking tools, and email enable effective communication among neuroscientists.

NEUROSCIENCE

Neuroscience is a field of investigation devoted to the observation of how the brain and the other components of the nervous system function, with a particular focus on their role in the cognitive, affective, and behavioral performance of an organism. Neuroscience includes many domains of research, being interrelated with such disciplines as neurobiology, neuroanatomy, neuropathology,

Areas of Research

Neuroscience's focus can be further categorized by taking into account the anatomical and functional attributes of the nervous system. Regarding anatomy, the central nervous system (CNS) and peripheral nervous system (PNS) are researched. The CNS consists of the brain and the spinal cord; the PNS is composed of spinal and cranial nerves. Considering functional classification, the autonomic nervous system is responsible for the nerves in smooth muscles and glands, whereas the somatic nervous system provides nerves to skeletal muscles and the skin. Studies of the brain do not only concentrate on how a given part of the brain is responsible for motor functions, visual perception, and verbal performance but also how the injury or illness of a given brain part leads to disorders and affects the whole performance of a person. Moreover, neuroscientists focus on the brain's performance during activities and sleep. Neuroscience also concentrates on mental illnesses (such as, e.g., schizophrenia), phobias (e.g., agoraphobia, or fear of open spaces; aviophobia, or fear of flying; arachnophobia, or fear of spiders) as well as eating or sleep disorders. Another field of interest is the notion of sensory systems, investigating the organs of taste and taste transduction. As for the visual sense, neuroscientists research, among other topics, how images are formed by the eye, how motion is perceived, and how colors and images are detected. Neuroscientists are also interested in the change of the nervous system throughout one's life, studying the changes in the brain from the embryonic stage until old age. Thus, neuroscience facilitates the understanding of the human body, its ordinary performance as well as different types of dysfunctions or problems. Neuroscience can also be subcategorized by taking into account the level of analysis in use. There are different specialties within neuroscience, such as molecular neuroscience, systems neuroscience, computational neuroscience, behavioral neuroscience, cognitive neuroscience, and clinical neuroscience, to mention a few. Starting from the smallest elements of investigation, such as molecules or cells, *molecular neuroscience* as well as *cellular neuroscience* can be researched; a more macro approach is represented by *systems neuroscience*, which benefits from the achievements in neuroscience and

biology in studying neurons and their circuits and networks. *Computational neuroscience* focuses on constructing the models of brain function by using mathematics and computer studies. Neuroscience also contributes to the understanding of individuals' behavior, cognition, and emotions. When one wants to observe the behavior of human beings, *behavioral neuroscience* offers the investigation of gender or age differences. *Cognitive neuroscience*, on the other hand, facilitates the understanding of people's emotions or reactions. To narrow the scope of the research, neuroscience may concentrate on one aspect of one's behavior. For example, *affective neuroscience* focuses on the neural side of emotions, by studying the behavior of human beings or animals. *Developmental neuroscience* observes the changes in the brain and its growth, whereas *clinical neuroscience* focuses on neurological disorders and brain injuries and involves treating patients in hospitals and other specialized health centers.

Regarding the scope of investigation, neuroscience focuses on both the smallest and the biggest entities responsible for the functioning of the brain and the nervous system. Taking into account the smallest elements of neuroscientific investigations, the neuron is one of the key components studied by neuroscientists. This type of cell is found in the bodies of human beings and almost all types of animals. Neurons possess the zones called synapses that enable them to communicate with other neurons, mainly by chemical neurotransmission. The studies on neurons and their electrophysiological properties facilitate the understanding how such complex systems operate and the way human bodies function. In addition, neuroscientists are interested in, for example, the parts of the brain, such as the cerebrum, consisting of frontal lobes, parietal lobes, temporal lobes, and their role in how a body functions. At the same time, neuroscientists look at the body as a whole since the health problems characterized by neurological symptoms, such as headaches, speech disorders, and numbness of limbs may have their roots in some dysfunction of one of the bodily organs, not necessarily the brain.

Neuroscience and Other Disciplines

The vast scope of issues and the complexity of the brain and the nervous system make neuroscience

crucial for the development of other areas of study. Taking into account the role of communication and language for human beings in social contacts, neuroscience is crucial to the sphere of language. The interrelation of these fields of investigation has given rise to *neurolinguistics*, which focuses on such topics as linguistic preferences, bilingualism, and multilingualism. In addition, neurolinguists are interested in language-related neurological problems or illnesses, such as aphasia, dementia, dyslexia, and dysgraphia. Thus, the research is devoted to how languages are perceived, learned, and used, taking into account language understood in different ways. Thus, not only languages understood as the official languages shared by a large group of people (e.g., a nation) are investigated, but also professional discourses, dialects, and genres are studied. Consequently, the learning and teaching processes are researched by studying, for example, mother tongues, foreign languages, and minority languages. Taking into account the example of linguistics, the subdisciplines of neuroscience can be characterized by interrelations. Since a language does not exist in a vacuum but is an important element of other fields of research, neurolinguists may work together with neuromarketers, studying the neurodimension of linguistic branding or advertising, or researching how customers select products and services on the basis of the offered linguistic information or stimulus. In the case of *neuromanagement*, this field may provide knowledge of how language is studied in professional settings, with observation of the role of language policies and the type of management as such. Moreover, *neuroethics* analyzes how the used linguistic tools strengthen, say, the policy of corporate social responsibility. Neuroscience is also crucial for education, leading to the appearance of such disciplines as *neuroeducation* or educational neuroscience. In addition, such subfields as neuropedagogy, neurodidactics, neuroteaching, and neurolearning can be named. Moreover, those working in education can also benefit from the knowledge of other domains relying on neurostudies, such as neurolinguistics, neuromanagement, neuroeconomics, cognitive neuroscience, cultural neuroscience, social neuroscience, and neuropsychology. Another field of investigation, *neurobiology*, is more focused on the biological

aspects of brain performance and thus, in the case of linguistics, can provide additional information on the part of the brain and its role in linguistic performance. Neurobiology can also be useful in clinical and therapeutical treatments, helping people recovering from brain injuries or facing some health problems to regain their speaking abilities. Neuroscience can also be used in studying the nature of art and beauty. *Neuro-aesthetics* focuses on studying aesthetics by means of neuroscientific methods and approaches. Such investigations offering information how art and music are perceived and experienced are of importance not only for neuroscientists but also for artists, historians, and those participating in the art life.

Neuroscientific Research

The issues of interest in neuroscience can be studied through the *4R model of neuroscientific research*, which focuses on the societal, behavioral, cognitive, and biological dimensions of one's performance and thus can be grouped around four main fields of investigation: Recognition, Reaction, Remembering, and Relations. The model, created by the author of this entry, is presented in Figure 1.

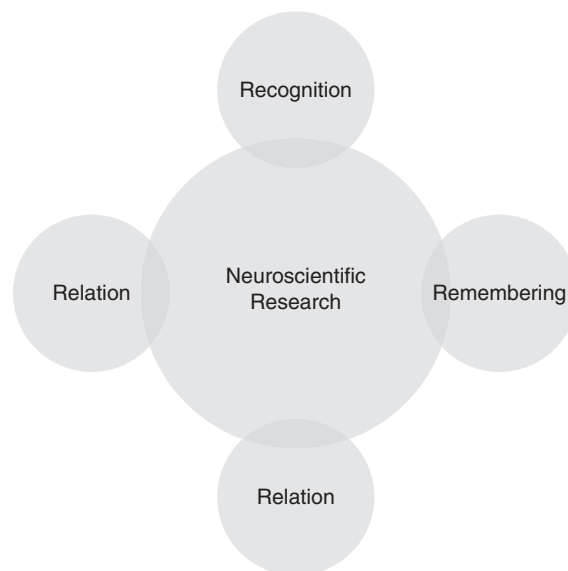


Figure 1 4R Model of Neuroscientific Research

Source: Magdalena Bielenia-Grajewska.

Recognition

Neuroscientists are interested how individuals recognize and perceive reality. One of the aims is to study how the brain and the nervous system take part in various behavioral and cognitive activities. Topics of study include the features of entities that make them attract the attention of subjects. Among the types of research available in the literature are those that focus on the role of colors, words, and images in capturing the attention of individuals and enabling these features to be recognized and categorized. Taking the linguistic type of study into account, research may focus on the role of metaphors in describing novel and complex concepts and recognizing their meaning. In addition, recognition can be studied by taking into account cultural differences in recognizing different concepts. For example, a figure from a fairy tale may be unknown in a culture that is not acquainted with a particular story. Thus, using this cultural concept in international advertising may be ineffective insofar as it requires familiarity with a culture-specific motif.

Remembering

Neuroscience offers explanations for how people perceive, acquire, store, and later apply knowledge. For example, the role of presenting data is studied, and its implications for understanding and using novel concepts are given close consideration. As far as neurolinguistic studies are concerned, researchers are interested how literal and nonliteral expressions are understood by individuals and which of them are more easily remembered. In addition, the neuroscientists study the role of repetition in learning new things as well as storing in memory the information already acquired. Another field of investigation is the nature of processes in short-term and long-term memory. Scientists are also interested in the role of representation in storing and using novel data. For example, they investigate the role of verbal and nonverbal stimuli in understanding new information. Research may involve studies of how pictures and sounds may facilitate the understanding and memorizing of information. Such findings are applicable, for example, in education, guiding teachers in how to choose materials to make

learning processes more efficient; and in neuro-branding, guiding marketers in how to make a brand remain longer in the minds of customers. Those interested in memory also take into account how different factors determine the possibilities of storing knowledge. For example, such factors as age and gender may be used to study people's differences in perceiving and memorizing data, thus forming a basis for the provision of diversified materials by taking into account the differing abilities of users. The study of people with memorizing deficits offers information as to which methods may prove most useful in teaching individuals having problems with perception and memorization, such as those with dementia or Alzheimer's disease.

Reaction

Reactions may be studied by observing subjects' responses to a presented stimulus. Such experiments involve the presentation of verbal and nonverbal cues and observing and measuring the reply to the incentives. These types of experiments are especially popular in marketing since they allow to study how people perceive a product or a service and what reactions follow the exposure to a given brand. Neuroscience is also interested in what makes reactions different from person to person, and the implications for evoking responses, both voluntary and involuntary. In addition, reactions may be studied by taking into account one's deficits as well as absent or exaggerated response to a stimulus. Also of interest are the ways in which some substances influence people's reactions. For example, neuroscientists investigate how alcohol, nicotine, medical substances, herbs, or narcotics may affect the performance of the brain and the nervous system. Such studies may develop the knowledge in the field of addictive substances to observe how certain substances may have deleterious effects on the brain and other organs. On another front, those interested in nutrition may be provided with more detailed information concerning proper diet and the influence of various foods on motor, cognitive, and behavioral functions of a human body. Proper nutrition becomes especially important for children, seniors, and people with illness since all have specific nutritional needs according to category.

Relation

Neuroscience facilitates the understanding of internal and external relations. Taking into account *actor–network theory*, neuroscience facilitates understanding about the role of human beings and nonliving entities in shaping reality. Neuroscientific tools offer information about the relation between different elements and their role in the overall performance of a given entity. Observing this feature in the linguistic domain, neuroscientists may investigate which sentences or words determine the efficient perception of a text. Another language-related issue may be studied by examining the role of literal and nonliteral elements in making a piece of information more easily understood by a diverse audience. Apart from the subject approach, neuroscience may also study human relations. By examining, for example, one's response to a given social stimulus, the role of interaction with other individuals may also be investigated. In addition, neuroscience facilitates the understanding of relations between neuroscience and other domains of knowledge—for example, neurolaw, neurolinguistics, or neuromanagement—which, at the same time, contribute to other domains of science and benefit from other disciplines.

Approaches and Tools Used in Neuroscience

Neuroscience relies on different methods and devices in conducting neuroscientific experiments. One possibility of categorizing them is to take into account the function they have. Thus, it is possible to discuss methodology used for scientific and clinical purposes. In addition, neuroscientific tools can be divided by taking into account the subjects. Thus, different devices are used for investigating animals and for investigating human beings. The approaches to conducting neuroscientific experiments can also be divided into invasive and noninvasive ones. Noninvasive techniques are used both for diagnosing and researching. One advanced apparatus is *functional magnetic resonance imaging* (fMRI); this neuroimaging technique relies on the phenomenon of magnetic resonance to observe brain performance. The subject lies still within an MRI scanner for about 60 to 90 minutes. The first stage takes about 6 to 15 minutes, and it is devoted

to making anatomical scans of the brain. After this phase, the subject is asked to perform some activities. Depending on the nature of an experiment (e.g., neurobranding or neurolinguistic studies), the activities may involve, for example, selecting an option, answering a question, and finding the difference between the presented options. During the execution of these tasks, the apparatus notices and measures the *blood oxygen level depending* signal that is responsible for showing which brain areas are active during a given activity. In the next stage of an experiment, the results are compared with the control scans done before the experiment itself. Neuroscientists also study facial expressions to perform noninvasive research. One such tool is *facial electromyography*, which provides information on one's emotions and feelings. Since one's face shows the emotional state, behavioral, cognitive, and affective notions may be studied by measuring the facial muscles, which are innervated by the facial nerves. For example, observing the zygomatic major muscle provides information on one's attitude to the presented verbal or nonverbal stimulus. This type of research may be used, for example, in advertising to check the reaction of the audience to a presented commercial. Another popular tool in neuroscience is *electroencephalography* (EEG). In this experiment, a set of electrodes are positioned at specific locations on the scalp of a subject. The experimenter observes the neural activities of the subject's brain when the person lies still and, for example, during exposure to light. This technique is noninvasive; however, EEG is also used as an invasive method in, for example, brain surgery—in that case intracranial EEG is used, and the electrodes are placed on the exposed part of the brain following craniotomy, where a surgical opening is made into the skull. Invasive techniques are used when it is not possible to diagnose a patient properly by using a noninvasive method. Another invasive technique, called *single-photon emission computed tomography* (SPECT), is used to determine the flow of blood to tissues and vessels. Used in neuro-oncology, SPECT offers information on the state of examined organs by using a radioactive substance and a camera. Intracranial blood vessels are studied by *angiography*, using a contrast agent and X-rays. Another dysfunction of blood vessels related to the improper velocity of blood flow in the brain can be studied

by *transcranial Doppler* (TCD) ultrasound. Moreover, neuroscience relies on psychological experiments and tools. Psychological tests are often administered in the first stage of research to help adjust and refine the neurological methods to be used. This option turns out to be especially useful in reducing the relatively high costs of conducting neuroscientific research.

Similarly, linguistic and economic types of research conducted prior to performing a neuroscientific experiment rely on such methods as questionnaires, interviews, and discourse analysis that precede, for example, fMRI. Neuroscientific experiments can also be categorized according to the different types of stimuli used by neuroscientists. Their selection depends on a number of factors, such as the age of subjects; for example, if young children take part in an experiment, audio or visual stimuli are used instead of written ones. Another factor may be the health condition of a subject. In addition, experiments involving non-native speakers of the language in which research is being conducted should take into account that language fluency may differ among the subjects. Using different neuroscientific methods also requires attention to ethics. Subjects taking part in neuroscientific testing should be informed of how the experiment will be conducted and how the information obtained in testing will then be used. In addition, people should be informed whether the tests are invasive, and, if so, what risks are involved.

Future of Neuroscience

It can be predicted that neuroscience will foster the development of studies that currently do not rely on neuroscientific advancements. Taking into account, for example, neuromanagement and neuroeconomics and the application of neuroscientific achievements in different domains of business studies, it is likely that other fields of economic investigation will benefit from neuroscientific knowledge. One reason for the growing interest in neuroscientific methods is their preciseness and the ability to study phenomena that cannot be observed by other means of investigation. In addition, neuroscientific methods eliminate the possibility of providing fake answers by respondents. Another possible field of development in

neuroscience is connected with achievements in other domains that have a bearing on neuroscientific investigation. Thus, the ongoing advances in MRI are leading in turn to advances in neuroscientific methods. The continuing development of neuroscience will lead to progress in at least two key areas. In health care, advanced neuroscience will be more effective in curing neurological diseases and brain injuries. In addition, neuroscience will contribute to the prevention of neurological problems and aid in slowing the progress of neurodegenerative disorders, such as Alzheimer's disease. Technical advances in neurological equipment may provide quicker, more effective, and more economical methods of diagnosing patients, introducing new methods that are noninvasive. Apart from the clinical perspective, advanced neuroscience offers better understanding of social life by studying behaviors and responses to different stimuli present in everyday life. Neuroscience may offer the answer how to make communication more effective and consequently make the health care discourse easier and more beneficial for patients, families, and professionals.

Magdalena Bielenia-Grajewska

See also Actor–Network Theory; Cognitive Science

Further Readings

- Bear, M. F., Connors, B. W., & Paradiso, M. A. (2007). *Neuroscience: Exploring the brain*. Williams & Wilkins.
- Bielenia-Grajewska, M. (2013). International neuromanagement. In D. Tsang, H. H. Kazeroony, & G. Ellis. (Eds.), *The Routledge companion to international management education*. Routledge.
- Bielenia-Grajewska, M. (2014). CSR online communication: The metaphorical dimension of CSR discourse in the food industry. In R. Tench, W. Sun, & B. Jones (Eds.), *Critical studies on corporate responsibility, governance and sustainability: Vol. 6. Communicating corporate social responsibility: Perspectives and practice*. Emerald Group, Ltd.
- Bielenia-Grajewska, M. (2015). Neuroscience and learning. In R. Gunstone (Ed.), *Encyclopedia of science education*. Springer. 698–704. https://doi.org/10.1007/978-94-007-2150-0_126
- Bielenia-Grajewska, M. (2019). Neuroscience and the food industry. In D. M. Kaplan (Ed.), *Encyclopedia of*

- food and agricultural ethics*. Springer. 1888–1892.
<https://doi.org/10.1007/978-94-007-6167-4-290-1>
- Helander, M. G., & Khalid, H. M. (2012). Affective engineering and design. In G. Salvendy (Ed.), *Handbook of human factors and ergonomics*. Wiley & Sons, 569–596.
- Kenning, P., Plassmann, H., & Ahlert, D. (2007). Applications of functional magnetic resonance imaging for market research. *Qualitative Market Research*, 10(2).
- Lee, N., & Chamberlain, L. (2007). Neuroimaging and psychophysiological measurement in organizational research: An agenda for research in organizational

- cognitive neuroscience. *Annals of the New York Academy of Sciences*, 1118, 18–42. <https://doi.org/10.1196/annals.1412.003>
- Siegel, A., & Sapru, H. N. (2015). *Essential neuroscience*. Williams & Wilkins.

NUMBER THEORY

See Mathematics, 19th Century;
Mathematics, 20th Century

P

PARACONSISTENCY

Paraconsistency is the study of logical systems with a nonexplosive negation such that a pair of contradictory formulas (with respect to such negation) does not necessarily imply triviality, discordant to what would be expected by the contemporary logical orthodoxy. From a purely logical point of view, the significance of paraconsistency relies on the meticulous distinction between the general notions of *contradictoriness* and of *triviality* of a theory—respectively, the fact that a given theory proves a proposition and its negation, and the fact that a given theory proves any proposition (in the language of its underlying logic). Aside from this simple rationale, the formal techniques and approaches that meet the latter definitional requirement are manifold. Furthermore, it is not solely the logical–mathematical properties of such systems that are open to debate. Rather, there are several foundational and philosophical questions worth studying, including the very question about the nature of the contradictions allowed by paraconsistentists. This entry aims to advance a brief account on some distinct approaches to paraconsistency, providing a panorama on the development of paraconsistent logic.

Paraconsistent Logic

From a structural standpoint, it can be said that paraconsistency is the property of a consequence relation $\vdash$ for which the principle *ex contradictione*

[*sequitur*] *quodlibet* (from a contradiction, anything [follows]) does not hold in general; that is, $A, \neg A \vdash B$ is not always the case (for a given negation $\neg$ and arbitrary propositions A and B). Accordingly, one can simply say that a logic is paraconsistent if and only if its consequence relation is not explosive.

Some other definitions of paraconsistent logic can be found in the literature. However, all of them can be shown equivalent with the latter when appropriate qualifications about the properties of the underlying inference relations are considered. Thus, a brief account on the development of such systems is appropriate.

The Beginnings (Modern Era)

The study of logical systems underlying possibly contradictory theories had risen about the same time that the constitution of the logicist and formalist schools in the philosophy of mathematics, at the end of the 19th century and the beginning of the 20th century. Indeed, already in 1910, Jan Łukasiewicz proposed a criticism on distinct formulations of Aristotle's views on contradiction. Following an equivalent route, in the same year, Nicolai A. Vasiliev advanced a kind of reasoning free from the laws of excluded middle and contradiction—called *imaginary logic* as an analogy with Nikolai Lobachevsky's *imaginary geometry* (a non-Euclidean geometry aiming to investigate the independence of the postulate of the parallels). Albeit endowing characteristics of a paraconsistent logic, namely criticizing the principle of noncontradiction, both works would mark

the birth of two other kinds of well-studied nonclassical logics—notably many-valued logics and dialectical logics, respectively.

Other works that pioneered the development of paraconsistent logic also adhere to a many-valued approach. The *logic of nonsense*, introduced by Sören Halldén in 1949, aimed at studying logical paradoxes by means of three-valued logical matrices, closely related to the *nonsense logic* introduced by Dmitrii A. Bochvar in 1938. An analogous approach was made in 1966 by Florencio G. Asenjo in his *Calculus of Antinomies*, which introduces a formal framework for studying antinomies by means of three-valued Kleene's truth tables for negation and conjunction, where the third truth-value is distinguished.

The Rise of Paraconsistent Logic as a Discipline

The contemporary growth of paraconsistent logic came about with translations of the first papers into English and an increase of interest in them. Some advancements that had marked the emergence of paraconsistent logic as a discipline, pushing forward novel results, include the works of Stanisław Jaśkowski and Newton da Costa, who independently developed comprehensive systems explicitly focusing on avoiding triviality by restraining the explosive character of contradictions, precisely the major characteristic of a paraconsistent logic—a term coined by Francisco Miró Quesada in a personal letter to da Costa in 1975, and then made public at the Third Latin America Conference on Mathematical Logic in 1976. The term *paraconsistency* is formed by the prefix *para-*, meaning “further than,” “beyond,” or even “similar to,” or “quasi”; and *consistency*, meaning the property of a system that is nontrivial or noncontradictory—two classically equivalent notions whose separation lies at the heart of the paraconsistentist agenda.

According to Jaśkowski, there are three main conditions that a contradictory yet nontrivial theory must satisfy: (1) it must be nonexplosive; (2) it should be rich enough to enable practical inference; and (3) it should have an intuitive justification. These latter two conditions, less formal than the first one, capture the biggest challenge to a paraconsistentist—namely the necessity to show that a certain account of logical consequence is somehow concerned with actual situations of reasoning.

Jaśkowski's motivation for advancing in 1948 the so-called *discussive logic* was a puzzling situation posed by Jan Łukasiewicz (of whom he was a disciple): Which logic applies in the situation where one has to defend some judgment A, also considering its negation for the sake of the argument? The strategy followed by Jaśkowski was to avoid the combination of conflicting information by precluding the rule of adjunction, making room for A and $\neg A$ without entailing the conjunction of both ($A \wedge \neg A$) since the classic explosion actually still holds in the form of $A \wedge \neg A \vdash B$. In terms of reasoning, it has a straightforward meaning: Each agent must still be individually consistent.

For da Costa, the focus is the development of systems that, on the one hand, could be strong enough to capture most of mathematics and, on the other hand, could circumvent some of the well-known paradoxes that historically had marked studies in logic. Da Costa's main intuition in the so-called *C-systems*, advanced in his 1963 paper “On the Theory of Inconsistent Formal Systems,” is that it is possible to differentiate inconsistent sentences from consistent ones or, in his own words, from those with a *well-behavior*—the latter being a sufficient requisite to guarantee the explosive character of a given formula.

Several other works in the literature arguably contributed to the development of paraconsistent logic as a mature discipline, most of them serving as a formal basis for particular research programs around which distinguished schools of paraconsistency have been organized. A brief account of those works is given in the next section from the standpoint of their respective approaches to paraconsistency.

Main Approaches to Paraconsistency

Preservationism

Out of the first works of Raymond Jennings and Peter Schotch on modal semantics, the early 1980s saw the organization of the preservationist school of paraconsistency, focused on advancing formal tools and intuitive justifications to capture the reasoning of human beings when faced with inconsistent data. Some of the core motivations include dealing with the concepts of belief and obligation, as well as C. I. Lewis's notion of *strict*

implication. Roughly speaking (apart from technical details and philosophical questions), given an inconsistent collection of sentences (in an already defined logic, usually classical logic), one should not try to reason about that collection as a whole but rather to focus on internally consistent subsets of premises. For introductory readings on the subject, see Schotch et al. (2009).

Adaptive Logics

A logic is said to be adaptive if it adapts itself to the specific premises to which it is applied. Developed out of the early work of Diderik Batens, adaptive reasoning recognizes some so-called abnormalities to develop formal strategies to deal with them: For instance, an abnormality might be an inconsistency (inconsistency-adaptive logics), or it might be an inductive inference, and a strategy might be to exclude a line of a proof (by marking it), or to change an inference rule. Thus, the paraconsistent character of the abnormality is not the main focus. Rather, the intuitive idea to be captured is that human reasoning can be better understood as endowed with many dynamic consequence relations, inconsistency-tolerant reasoning being one of them. Some papers on the subject can be found in a 2000 work by Batens and colleagues.

Relevant Logics

Mainly concerned in avoiding the paradoxes of material and strict implication, *relevant logics* are substructural logics that demand a meaningful connection between the premises and the conclusion of an argument. This strategy induces a paraconsistent character in the resulting deductions since A and $\neg A$, as premises, do not necessarily have a meaningful connection with an arbitrary conclusion B . Relevant logic evolved in the 1960s out of some early works by Alan Anderson and Nuel Belnap, themselves extending ideas of Wilhelm Ackermann, Alonzo Church, and others. For a comprehensive technical introduction, see Anderson et al. (1992).

Dialetheism

Dialetheia is a neologism formed by the Greek prefix *di[a]-*, meaning “in opposite or different directions”; and *aletheia*, a word for “truth.”

Accordingly, a *dialetheia* is a sentence that is both true and false. Developed out of the original works of Richard Routley and Graham Priest in the 1970s, mostly motivated by the advancement of tools to deal with the liar paradox and set theoretic antinomies, dialetheism is the ontological view that there are *dialetheia*, that is, the view that some contradictions are true. Therefore, as expected, the preferred logic for a dialetheist should be paraconsistent, although not all paraconsistentists are compelled to be dialetheists. For a comprehensive philosophical introduction, see Priest (1995).

Logics of Formal Inconsistency (LFIs)

A key idea here is that there are situations in which contradictions can, at least temporarily, be admissible if their behavior can be somehow controlled, as da Costa has it. Further generalizing and extending the advancements of the latter, the LFIs (as it would be christened by Walter Carnielli and João Marcos in early 2000s) are a family of logics that can encode [in]consistency within their object language, allowing an explicit separation between contradictoriness from inconsistency, inconsistency from triviality, consistency from noncontradictoriness, and nontriviality from consistency. For a state-of-the-art account of the subject, see Carnielli and Coniglio (2016).

Other Approaches

There are many ways a logic can be contradictory yet nontrivial. Jaśkowski’s discussive logic, for instance, not only followed a *nonadjunctivist* approach but also endowed a *modal* character—in fact being a fragment of the well-known modal logic S5. Another significant program is the *many-valued* one, for which the first contemporary comprehensive system with an explicit paraconsistent agenda is Asenjo’s calculus of antinomies, the same strategy followed by Priest’s *logic of paradox*, advanced in first works on dialetheism.

The fact is that there is no unique way to divide the advancements made in the literature into specific approaches, since they intersect and complement each other. Paraconsistency, as a property of logical systems, is negatively defined—it encompasses every system where *ex contradictione*

quodlibet does not hold in general. The division of those systems into distinct schools of paraconsistency, highlighting particular properties and motivations, can thus be understood solely as a pedagogical tool to introduce the rich and fruitful plurality of paraconsistent logics.

In addition to the cornerstone logic-philosophical debate, the study of paraconsistency from the perspective of finite models of arithmetic as well as the applications of paraconsistent logic in some computational areas have provided a new dimension to the ongoing debate. The fact is that since the first works in the area, paraconsistency has turned out to be a remarkably fertile research field that provides us with new ways to deal with contradictory yet nontrivial scenarios, including inconsistent theories, paradoxes, dialectics, ontology, belief dynamics, and many more.

Rafael R. Testa

See also Deduction; Logic, Formal and Informal; Paradoxes; Rationality; Reasoning

Further Readings

- Anderson, A. R., Belnap, N. D., & Dunn, J. M. (1992). *Entailment: The logic of relevance and necessity*. Princeton University Press.
- Asenjo, F. G. (1966). A calculus of antinomies. *Notre Dame Journal of Formal Logic*, 7(1), 103–105. <https://doi.org/10.1305/ndjfl/1093958482>
- Batens, D., Mortensen, C., Priest, G., & van Bendegem, J.-P. (Eds.). (2000). *Frontiers of paraconsistent logic*. Research Studies Press.
- Beziau, J.-Y., Carnielli, W. A., & Gabbay, D. M. (Eds.). (2007). Handbook of paraconsistency. In *Studies in logic* (Vol. 9). College Publications.
- Carnielli, W. A., & Coniglio, M. E. (2016). *Paraconsistent logic: Consistency, contradiction and negation*. Springer.
- Carnielli, W. A., Coniglio, M. E., & D'Ottaviano, I. M. L. (Eds.). (2002). Paraconsistency: The logical way to the inconsistent. *Lecture Notes in Pure and Applied Mathematics* (Vol. 228, pp. 1–94). Marcel Dekker.
- da Costa, N. C. A. (1974). On the theory of inconsistent formal systems. *Notre Dame Journal of Formal Logic*, 15(4), 497–510.
- Priest, G. (2002). *Beyond the limits of thought* (2nd ed.). Clarendon Press.
- Priest, G., Tanaka, K., & Weber, Z. (2018). Paraconsistent logic. In E. N. Zalta (Ed.), *The Stanford encyclopedia of philosophy*. <https://plato.stanford.edu/archives/sum2018/entries/logic-paraconsistent/>
- Schotch, P., Brown, B., & Jennings, R. (Eds.). (2009). *On preserving: Essays on preservationism and paraconsistent logic*. University of Toronto Press.
- Tanaka, K., Berto, F., Mares, E., & Paoli, F. (Eds.). (2013). Paraconsistency: Logic and applications. In *Logic, epistemology, and the unity of science* (Vol. 26). Springer.
- Weber, Z. (2022). Paraconsistent logic. In *The Internet Encyclopedia of Philosophy*. <https://iep.utm.edu/para-log/>

PARADIGM

The term *paradigm* comes from the Latin word *paradigma*, which, in turn, comes from the ancient Greek word *paradeigma* (παράδειγμα: “παρα,” beside, near; “δειγμα,” sample, what is shown). *Paradigm* and cognate expressions, such as “change of paradigm” or “paradigm shift,” are used in different contexts, both scientific and academic, as well as in everyday life.

Until the appearance of Thomas S. Kuhn’s work, the word *paradigm* was used primarily in two senses: (1) in rhetoric, as an example or case of something that serves as a model or pattern for other cases of the same that may be copied; as a very clear and typical example of something, as a type-example; (2) in grammar, as an example of a conjugation or declension showing a word in all its inflectional forms, creating a pattern of conjugating or declining, where other words of the same type conjugate or decline in an analogous way.

But it is only since Kuhn’s work that the term *paradigm* and related expressions have acquired the diffusion and widespread use attained today. And although Kuhn introduces the term in the context of theoretical reflection on science, it is incorporated into colloquial language with an even more encompassing meaning, in the sense of a philosophical or theoretical framework of any kind, or in the more general sense of a perspective, a position, a view, a way of looking at something or regarding a situation or topic.

In fact, we are faced with a paradoxical situation: Although many people are familiar with the term *paradigm* and continue to use it in one sense

or another, Kuhn himself, owing to the term's ambiguity, had ceased to do so—although without renouncing the concepts that led him to its introduction. Moreover, such concepts continue to be of fundamental importance in understanding the nature, functioning, and development of science.

This entry focuses on these Kuhnian concepts of paradigm. After a brief introduction to Kuhn's paradigms, it will first place the Kuhnian conception in the context of 20th-century philosophy of science. Kuhn's concepts of paradigm will then be outlined. Next, it will be seen how these concepts explain the pattern or regularity that Kuhn identifies in the development of science. It will continue with the identification of some relations of Kuhn's proposal with classical philosophy of science. The entry concludes by pointing out some further developments of Kuhn's paradigms.

Introduction to Kuhn's Paradigms

Although the term *paradigm* is used by Kuhn for the first time in the 1959 text "The Essential Tension: Tradition and Innovation in Scientific Research," in the sense of what he would later call *exemplar*, it is with his 1962 work, *The Structure of Scientific Revolutions* (SSR), that it became customary in philosophical reflection and in historical analyses of science.

In this work, which seeks to promote a more authentically historical history of science, contra *presentism*, *whiggism*, or the *Whig interpretation of history*, Kuhn develops a theory of the history of science that opposes the received image of scientific development, which considers it to be cumulative, continuous, and linear.

This theory of the history of science has strong implications for the philosophy of science, where the philosophical importance of historiographical reform is manifested.

For, insofar as the philosophy of science is based on the history of science and forms its image of the practices and rationality of science from it, if the history of science is wrong, so will be the philosophy of science *modeled* on it.

Thus, SSR also promotes a revision of philosophers' image of science, in particular of scientific change, intending to bring philosophy of science in tune with what scientists really do, and have done, throughout history, thereby forcing us to

rethink the concepts of rationality, progress, and scientific development.

In time, SSR ended up becoming a true best seller and has marked our conception of science ever since.

The Kuhnian Conception in the Context of 20th-Century Philosophy of Science

The prevalent image of philosophers of science by the mid-20th century was that of the so-called classical conception (or *received view*) of theories and some related ideas on science around topics such as concept formation, hypothesis testing, and scientific explanation. In its most general sense, the classical conception of theories can be characterized as explicating the concept of a scientific theory as a set of statements, sentences, or propositions deductively or axiomatically organized.

By contrast, in a more detailed formulation, such as that outlined by Rudolf Carnap, we can distinguish three general aspects in the explication of the concept of theory. The first one refers to the (more) *theoretical* (or *formal*) part that is constituted by the formal axiomatic system or *calculus* (symbolized by *T*)—which contain only descriptive theoretical terms. The second one corresponds to the (more) *empirical* or *testing* part that is given by pure observational statements—which contain only descriptive observational terms. The third one establishes the relationship between *theory* and *experience* through linguistic means, the so-called correspondence rules (symbolized by *C*)—which connect theoretical terms with observational terms. Thus, *the* theory, or interpreted calculus, consists of the conjunction of all the axioms and all the correspondence rules *T* and *C*.

Beginning in the 1950s, this view of scientific theories has been subject to criticism, of which there are mainly two kinds: (a) criticisms of certain aspects of the classical conception (e.g., of the distinction between theoretical and observational terms) and (b) a global criticism, which attacks mainly the bases of the conception, proposing an alternative view on science. The second kind of criticisms of the classical conception came mainly from people interested in the history of science, once referred to as *new philosophers*, giving rise to what would be called the *historical turn* in the

philosophy of science, of which Kuhn is its best-known representative. A new conception about the nature and the synchronic structure of scientific theories (without this being implied in a strict sense and without its being systematically developed) underlies the majority of diachronic studies and analyses, typical of this historicist phase, which is supposed to be closer to scientific practice, as history presents it to us.

According to the new conception, scientific theories, which the new philosophers refer to with different terms (paradigms or *disciplinary matrixes* for Kuhn, *research programs* for Imre Lakatos, *research traditions* for Larry Laudan), in order to avoid being confused with the classical conception, are not sentences or sentence sequences, and in a proper sense they cannot be described as true or false (although true or false empirical claims are certainly made with them), but they are highly complex and ductile entities, susceptible of evolving in time without losing their identity.

But Kuhn's proposal should also be placed within another of the *turns* that have taken place in analytic philosophy in general and philosophy of science in particular, namely the so-called *pragmatic turn*. This turn in philosophy of science arises from the rejection of what Wolfgang Stegmüller calls *the third dogma of empiricism*, namely the conviction that for the explication of all epistemologically relevant fundamental aspects of science the instruments of logic (syntactic and semantic tools) are sufficient.

This led to the acceptance that in the philosophical investigation of science, not only the syntactic and semantic aspects of language but also the pragmatic ones must be taken into consideration.

It is often said that the changes that occurred in the philosophy of science during the 1960s produced a real revolution in the field. If, however, we take into account the multiplicity and variety of positions held by philosophers of science in the first half of the 20th century, it would perhaps be better to characterize these changes as a recovery or deepening of the problems addressed and of the solutions previously advanced, for example, by logical empiricists such as Otto Neurath, Edgar Zilsel, Philip Frank, or even Rudolf Carnap, or by people outside this philosophical movement, in particular Ludwik Fleck.

Nevertheless, the incidence of the new philosophers was decisive in this resurgence in the 1960s: The consideration of the historical or historicist perspective that generally characterizes them definitely marks the development of later metascientific reflection.

Kuhn's Concept(s) of Paradigm

Since the appearance of the first edition of SSR in 1962, the notion of *paradigm*, central to Kuhn's conception of science, has been critiqued for its vagueness and ambiguity. One commentator went so far as to point out 21 different senses of this term, while acknowledging that not all of them are inconsistent with each other.

Kuhn took this criticism seriously. In 1969, he wrote three works—the book chapters “Second Thoughts on Paradigms,” “Reflection on My Critics,” and a postscript for the second edition of SSR—with the aim of clarifying some points of view developed in SSR, including his conception of paradigms. In them, he claims to have been using the term *paradigm* basically in two different senses: (1) as the global set of commitments shared by the members of a given scientific community and (2) as concrete solutions to problems.

To avoid misunderstandings, he proposes to replace the term *paradigm* by *disciplinary matrix* to refer to the first, *global* sense of the term, and by *exemplars* to refer to the second, *original* sense of the term. The following are the elements that constitute a disciplinary matrix:

1. *Symbolic generalizations*, which are law schemes not discussed by scientists, formalized or easily formalized, and which act partly as definitions and partly as genuine laws, establishing the most general relationships between the entities that populate the field under investigation (and corresponding to Kuhnian explication of the concept of fundamental law).
2. *Ontological or heuristic models*, which manifest the ontological or metaphysical convictions as to what there is and what its fundamental characteristics are, and which give the group its preferred or permissible analogies and metaphors and make it possible to visualize and make its behavior more comprehensible.

3. *Methodological values*, which may be shared with other disciplinary matrixes, such as accuracy—exact agreement, or not exceeding a certain margin of error, with the results of existing experiments and observations; consistency—internally and with other currently accepted theories; scope—extending far beyond the particular observations; simplicity—bringing order to phenomena that otherwise would be individually isolated and, as a set, confused; and fruitfulness—disclosing new phenomena or previously unnoted relationships among those already known.
4. *Shared examples or exemplars*, which Kuhn considers the most original and least understood aspect of his book, and which constitutes the original meaning of the term *paradigm*, as introduced in his (Kuhn, 1959) in analogy with its use in language teaching. They are concrete solutions that successfully solve problems posed by the paradigm-disciplinary matrix, adapting the symbolic generalizations and obtaining the specific symbolic forms required by particular problems; and which show scientists in a nondiscursive way what entities populate the universe of research, what questions can be asked, what are the admissible answers, and what are the methods for testing them. Shared examples or exemplars constitute sense 2 of *paradigm*, distinguished above.

The Development of Science

The nature and function of paradigms, both in the sense of disciplinary matrix and in the sense of exemplar, explain the pattern or regularity that Kuhn identifies in the development of the mature sciences, that is, the successive phases, stages or periods—after an initial one of preparadigm (prenormal or preconsensus) science—of normal science, crisis and extraordinary science, new normal science, new crisis and extraordinary science . . . (what constitutes “the structure of scientific development” in general), and the successive transition from one paradigm to another through a scientific revolution (what constitutes “the structure of scientific revolutions” in particular).

The preparadigm (or, perhaps better, prenormal or preconsensus) period is characterized by the

existence of competing different schools and subschools working in different directions, without common commitments.

After a paradigm-disciplinary matrix achieves the consensus of the scientific community, a broad avenue for research opens up, in the form of closely related problem-solving or, insofar as they are supposed to have an assured solution within the accepted paradigm-disciplinary matrix, *puzzle-solving*, which scientists carry out under its guidance over a long period, called *normal science*. Scientists recognize the problems posed by the paradigm-disciplinary matrix as similar to the shared examples or exemplars and solve them in a manner similar to the shared examples. Through this practice, the paradigm-disciplinary matrix achieves greater precision and articulation within itself and with nature; that is, it also broadens its domain of application.

During this period of normal science, scientists work with the conviction that this or that problem will have a solution within the conceptual framework of the paradigm-disciplinary matrix, proposing, in a hypothetical manner, that a certain modification of the symbolic generalization—though not obtained by deduction (just) from it—will do the job. If the proposal of a specific symbolic form is successful, the applicability of the paradigm-disciplinary matrix to reality is extended, affirming it in its fertility. In the case of its refutation, the discredited one is, according to Kuhn, the scientist himself—who fails to propose the appropriate specific form of the symbolic generalization that would solve the problem posed, one that would satisfactorily fit the data obtained—and not the paradigm-disciplinary matrix.

Faced with a negative test, the symbolic generalization of the paradigm-disciplinary matrix is, or can be, always safeguarded by modifying the nonnuclear elements, in other words the specific form proposed by the scientist to solve the *puzzle* posed by the paradigm-disciplinary matrix and whose resolution would be assured.

However, when frustrations accumulate while trying to solve *problems* that should be solved, these go from being the driving force of the development of the paradigm-disciplinary matrix to being perceived as *anomalies* whose existence compromises the usefulness of the paradigm-disciplinary matrix for research. A period of *crisis* and *extraordinary science* begins. A small group of

researchers starts to work outside the accepted paradigm-disciplinary matrix from new perspectives that are incompatible with the previous ones, until a new paradigm-disciplinary matrix is established that succeeds in problem areas that the scientific community considers important, and promises to solve others, some of which were not even on the agenda of the previous one, and giving up others, which are now no longer considered legitimate. The bulk of the scientific community begins to abandon a paradigm-disciplinary matrix exhausted in its heuristics, to adopt the one that allows it to leave behind the feeling of futility of its own work, consummating a *scientific revolution* and giving rise to a new period of normal science.

The rupture present in the shift from the old paradigm-disciplinary matrix to the new one thus entails not only gain but also loss. Although incommensurable—meaning by this that there is no common or neutral basis with which to measure both paradigm-disciplinary matrixes or a common or neutral (observational) language that allows the intertranslatability of both paradigm-disciplinary matrixes without waste or loss—the process of abandoning one paradigm-disciplinary matrix and simultaneously accepting another is not irrational, as Kuhn's critics thought. The choice between paradigm-disciplinary matrixes is not resolved by the application of norms or rules based only on logic (internal coherence) or experiment (external coherence). This does not imply, however, that there are no good reasons guiding such a choice, rather only that these reasons (among which are the aforementioned simplicity, accuracy, coherence, scope, and the ability to generate fruitful research) function as values or criteria shared by scientists but are capable of being applied differently by different researchers—a rationality of another type (practical), different from the one traditionally proposed (logical or theoretical), but as far from arbitrariness as the latter: less precise, debatable, with risks in the choice that the scientific community diminishes by distributing the danger among its members, until time shows with its results the rightness of the bet.

Kuhn's emphasis that the scientific community is inseparable from the theoretical and empirical elements of the paradigm-disciplinary matrix clearly differentiates his conception of science from the traditional ones. He will go so far as to say—in

a “circular, but not vicious” way—that a paradigm is what a scientific community shares, while a scientific community is one that shares a paradigm. There are several reasons for him to introduce this notion. On the one hand, the historian of science visualizes the changes in theories (paradigm-disciplinary matrixes) as a change in the beliefs of the only ones with the authority to decide them, the community of experts, in a context in which it was shown that there were no crucial facts that forced the discarding of one theory (paradigm-disciplinary matrix) and the adoption of another, nor a completely common language to guide the discussion. On the other hand, the existence of normal science means that the development of the paradigm-disciplinary matrix is not due to any isolated scientist but to the joint effort of a group of researchers that makes it advance when they solve under its guidance the innumerable problems it poses.

In addition, Kuhn proposes to abandon the *teleological notion of progress toward truth*, according to which changes in the paradigm-disciplinary matrix bring scientists ever closer to the truth, preferring instead to speak of a development—analogue to that proposed by the theory of evolution with respect to species—that can be defined *from* its previous stages, as opposed to a process of evolution *toward* something.

Kuhn's Relations With Classical Philosophy of Science

As regards his relations with classical philosophy of science, Kuhn—who expected to find in Popperians his best allies—tries to show how his ideas follow those of Karl Popper's, although in his own way.

The vehement rejection he suffered taught him that, although they coincided in some aspects, the Popperian community and Popper himself would not forgive him for the pragmatic, especially psychological and sociological, aspects of his proposal. From then on, he would be read as somebody who ascribes an irrational behavior to scientists instead of admitting the necessity (shown by his analyses) of modifying the concept of scientific rationality subscribed to until then.

The situation is equally paradoxical in relation to logical empiricism, which is supposed to be the adversary defeated by his work.

Apart from the stereotypes that turned it into the *straw man* everybody uses to revile it, this current in the philosophy of science presents a wide range of aspects and orientations which even justify the enthusiastic recommendation of Kuhn's SSR that Carnap wrote in his own hand on the back of the official letter of acceptance which he sent to Charles Morris.

There are clear affinities between Kuhn's proposal and that made by the logical empiricists Neurath and Zilsel, who carry out more *historical* and *sociological* analyses than the more *formalist* one represented by Carnap's work.

And there are affinities even with the latter, although Carnap did not consider the historical and social empirical studies of science as belonging to the philosophy of science, understood as the *logic* of science, but to the more inclusive category *theory* of science. Examples of such affinities include Carnap's acceptance of (1) the noncumulative conception of scientific development (positing a sort of incommensurability between different linguistic or conceptual frameworks), (2) the role of pragmatic considerations in the scientists' decision of using a linguistic framework, and (3) accepting or rejecting particular hypotheses (the latter also stressed by Neurath, Frank and Hempel's late work).

The fact that the *Postscript* to SSR written by Kuhn in 1969 was the last work published in the most ambitious publishing project coming from logical empiricism, the (*International*) *Encyclopedia of Unified Science*, constituted the perfect ending for an era, not because Kuhn finished with that tendency forever but because with him some interests that had begun in Vienna at the beginning of the century would find their way forward.

Further Developments

After the clarifying articles written in 1969, in which he replaces the term *paradigm* with the terms *disciplinary matrix* and *shared example* (or *exemplar*), Kuhn also stopped writing about disciplinary matrixes and instead wrote about *theories*. This doesn't mean, however, that Kuhn abandoned the *concept* of paradigm, in either of the two basic senses—the *original* sense of exemplar and the more overarching one of disciplinary matrix—but only the term.

Notwithstanding, that he uses the term *theory* doesn't mean that he accepted the classical concept of theory. According to his own explication of the concept of theory, it consists, among other things, of *symbolic generalizations*, in its (more) *theoretical* or *formal* part, *together with* examples of their function in use (*exemplars* or *paradigms* in the *original* sense), in its (more) *empirical* or *applicative* part. As for the link between the two parts, it is established by what Kuhn calls "special (or appropriate) versions" or "particular (or detailed) symbolic forms (or versions or expressions)," which acquire the symbolic generalizations in order to be applied to particular problems (situations, phenomena). And although Kuhn does not elaborate in detail what the relationship between symbolic generalizations and their particular forms is, he makes it very clear that this is not one of logical deduction.

Some historians and philosophers of biology have found the notion of *exemplar* fruitful for their analyses, either by holding that theories in the biological (and/or biomedical) sciences possess a particular structure distinct from that of physical theories as Darden (1991) and Schffner (1986) argue. or by considering that this is not the case, if analyzed within the framework of some version of the semantic conception of theories. Lorenzano (2007) and Schaffner (1993) provide such analyses.

In fact, Kuhn's concept of theory, which reformulates that of paradigm as a disciplinary matrix and contains the *original* sense of the former, finds an even more satisfactory explication in the framework of the semantic conception of theories, especially in that of the structuralist view, as Kuhn himself early recognized and continued to do so until the end of his days.

In the structuralist version of the semantic conception of theories that acknowledges the presence of irreducibly pragmatic and historically relative elements in the analysis of scientific theories as we see in Balzer et al. (1987), the Kuhnian notion of symbolic generalization is made more precise by means of the notion of *fundamental law/guiding principle*, that of exemplar through that of *paradigmatic application* and the relation between symbolic generalizations and the specific forms they adopt with the relation of *specialization*.

Whereas the overall synchronic structure of a theory is thus given by a *hierarchical (non-deductive) structure organized by the relation of specialization*

and the scientific development during the period of normal science is represented in it by the notion of *theory-evolution*.

And together with the (*usually partial*) *incommensurability* existing between theories separated by a scientific revolution, the structuralist view of theories also makes possible the representation of other inter-theoretical changes susceptible to be found in the history of science, such as the *emergence* of theories and the *reduction* between them.

Pablo Lorenzano

See also Abduction; Conceptual Analysis; Falsifiability; Hypothetico-Deductivism; Philosophy of Science; Received View of Theories; Scientific Revolutions

Further Readings

- Balzer, W., Moulines, C. U., & Sneed, J. D. (1987). *An architectonic for science. The structuralist program*. Reidel.
- Carnap, R. (1956). The methodological character of theoretical concepts. In H. Feigl & M. Scriven (Eds.), *Minnesota studies in the philosophy of science* (Vol. 1, pp. 38–76). University of Minnesota Press.
- Darden, L. (1991). *Theory change in science: Strategies from Mendelian genetics*. Oxford University Press.
- Kuhn, T. S. (1959). The essential tension: Tradition and innovation in scientific research. In C. W. Taylor (Ed.), *The third (1959) University of Utah research conference on the identification of scientific talent* (pp. 162–174). University of Utah Press.
- Kuhn, T. S. (1970a). Postscript—1969. In T. S. Kuhn (Ed.), *The structure of scientific revolutions* (2nd ed., pp. 174–210). University of Chicago Press. (Original work published 1969)
- Kuhn, T. S. (1970b). Reflections on my critics. In I. Lakatos & A. Musgrave (Eds.), *Criticism and the growth of knowledge* (pp. 231–278). Cambridge University Press. (Original work published 1969)
- Kuhn, T. S. (1970c). *The structure of scientific revolutions* (2nd ed.). University of Chicago Press. (Original work published 1962)
- Kuhn, T. S. (1977). Second thoughts on paradigms. In F. Suppe (Ed.), *The structure of scientific theories* (2nd ed., pp. 459–482). University of Illinois Press. (Original work published 1969)
- Kuhn, T. S. (1992). Introduction to presidential address. In D. Hull, M. Forbes, & K. Okruhlick (Eds.), *PSA 1992* (Vol. 2, pp. 3–5). Philosophy of Science Association. <https://doi.org/10.1086/psaprocbienmeetp.1992.2.192821>
- Kuhn, T. S. (2000). A discussion with Thomas S. Kuhn (with Aristides Baltas, Kostas Gavroglu, and Vassili Kindi). In J. Conant & J. Haugeland (Eds.), *The road since structure: Philosophical essays, 1970–1993, with an autobiographical interview* (pp. 255–323). The University of Chicago Press.
- Lakatos, I., & Musgrave, A. (Eds.). (1970). *Criticism and the growth of knowledge*. Cambridge University Press.
- Lorenzano, P. (2007). Exemplars, models and laws in classical genetics. In J. L. Falguera, M. C. Martínez, & J. M. Sagüillo (Eds.), *Current topics of logic and analytic philosophy* (pp. 89–102). University of Santiago de Compostela.
- Moulines, C. U. (2014). Intertheoretical relations and the dynamics of science. *Erkenntnis*, 79, 1505–1519. <https://doi.org/10.1007/s10670-013-9580-y>
- Reisch, G. A. (1991). Did Kuhn kill logical empiricism? *Philosophy of Science*, 58, 264–277. <https://doi.org/10.1086/289615>
- Schaffner, K. F. (1986). Exemplar reasoning about biological models and diseases: A relation between the philosophy of medicine and philosophy of science. *The Journal of Medicine and Philosophy*, 11, 63–80. <https://doi.org/10.1093/jmp/11.1.63>
- Schaffner, K. F. (1993). *Discovery and explanations in biology and medicine*. University of Chicago Press.

PARADOXES

Typically when we reason, we start from premises that we believe, proceed through inferences based on principles of logic that we could not doubt, and arrive to a conclusion that we accept. *Paradox* is the name we give to situations in which this *doesn't* happen. Paradoxes arise when an argument seems to show that something has gone awry in our reasoning, when the argument itself seems airtight and yet the conclusion seems absurd. This is different from making a mistake on one's math homework where an error in a calculation might lead one to mistakenly deduce $0 = 1$. But the difference is mainly one of degree; while it is usually straightforward to spot a dropped minus sign or division by 0 in math homework mistakes, paradoxes are not so easily resolved. Indeed, resolution of a paradox usually requires giving up a firmly held premise, adopting unfamiliar methods of reasoning, or accepting an intolerable conclusion. For this reason, the study of paradoxes has been

behind major upheavals of thought and scientific innovations, as the resolution of paradoxes required either new ways of reasoning or substantial revisions of commonly held beliefs.

This entry aims to explain what paradoxes are and the ways that they have been characterized and solved over the course of the history of philosophy. We will begin by describing some examples in detail, illustrating the ways that paradoxes can arise in considerations of infinity, vagueness, and self-reference. These examples will then serve as the raw materials for a subsequent discussion on the ways that paradoxes have been analyzed and understood. The study of paradoxes has been a major concern within philosophy for millennia, so the list of examples considered here and the discussion about them will inevitably be incomplete—we do not discuss, for example, moral paradoxes or paradoxes of rational action—but our aim will be to highlight the shared features of the paradoxes and the ways that people have tried to solve them. With these features in view, we will conclude with a discussion about *why* the resolution of paradoxes has, for various corners of philosophy and science, been an urgent priority and how these solutions have undergirded important conceptual advances in the relevant fields.

Examples of Paradoxes

Zeno's Paradox

Zeno's paradox is really a family of paradoxes that derive contradictory conclusions from the infinite divisibility of time and space. These paradoxes, originally due to Zeno the Greek in the 5th century B.C.E., have survived to the present day by being passed down secondhand through the work of other philosophers, principal among them being Aristotle. For the sake of brevity, we will focus on two versions of Zeno's paradox, which stem, in particular, from the fact that a finite interval of time can be broken up into an infinity of subintervals of positive length.

The first concerns motion from point A to point B. Suppose you begin a journey from point A to point B, which are at some positive distance apart from one another. Before you arrive to point B, you must arrive at the midpoint C_1 along the

path from A to B. Then, continuing along the path, before you reach B, you must also reach the midpoint C_2 along the path from C_1 to B. And this continues: Once you reach the point C_n , you must first reach the midpoint C_{n+1} along the path from C_n to B, for $n = 1, 2, 3, \dots$. Therefore, you have to perform infinitely many arrivals before you can arrive to point B. Infinitely many actions cannot be performed in finitely much time. Therefore, it is impossible to travel from point A to point B. Notice, moreover, that, in this argument, point A and point B were completely arbitrary locations at some distance apart, so this argument seems to show that going from anywhere to anywhere else leads to contradiction or, in other words, that motion is impossible.

A more vivid variant of this paradox is formulated in terms of a race between Achilles and the tortoise. Achilles is an athletic fellow and a swift runner, while the tortoise ambles about at a glacial pace. It should be clear that, even if the tortoise gets a head start, Achilles will at some point overtake the tortoise. Yet before Achilles overtakes the tortoise, he must advance to a location at which the distance between him and the tortoise has been halved. And then before he overtakes the tortoise, he must next reach a location at which *that* distance between him and the tortoise has been halved, and so on. Thus, it appears that Achilles, no matter how quickly he sprints, can never catch up to the tortoise.

Sorites Paradox

A second paradox, also of ancient origin, is called the *Sorites paradox*. This paradox concerns vagueness. One version of the argument runs as follows. If one removes a single grain of sand from a heap of sand, it is still a heap of sand. A single grain of sand, by itself, is certainly not a heap of sand. Now suppose we are given a heap of sand. It will contain some finite number of grains of sand. For concreteness, let's assume it consists of 14,379 grains of sand. Now suppose that, one by one, we remove the grains of sand from the heap. We perform the operation of removing a single grain 14,378 times, until we arrive at a single grain of sand left. So we arrive at something which is certainly not a heap of sand, but we started with a heap and, since the removal of a single grain

cannot turn a heap into a nonheap, it must be the case that the single grain of sand is a heap as well. This is a contradiction.

A similar puzzle, recorded by the ancient Greek historian and biographer Plutarch, concerns the fabled ship of Theseus, the fabled founder-king of Athens. Ships have parts that rot, break, and need to be replaced. After several years of constant sailing, Theseus's ship had undergone so much maintenance that literally every part been replaced and none of the physical parts of the original ship remained. Is the ship that resulted after years of sailing the same as the ship with which Theseus began? The fact that every piece of the latter ship differs from the former one seems to mean that it is a *completely different ship*. But one would not say that one has a new boat just because a part gets replaced or a plank gets mended. If the ship is different, then there must be some modification where the transformation took place, but none can be located in the ship's entire history of repair.

The Sorites paradox and related puzzles deal with vague concepts, whose characteristic feature is that they have borderline cases, where competent speakers do not know how to respond to a question about whether or not a concept applies. A single grain is not a heap, a story-high pile of sand certainly is, but somewhere in the middle, individuals will not necessarily know how to respond to the question of whether or not a certain quantity of sand constitutes a heap. The concepts considered in the Sorites paradox, moreover, satisfy a *principle of tolerance* whereby small changes to an object do not change whether or not the concept applies.

The Liar Paradox

The Liar paradox comes from trying to determine the truth or falsity of a sentence that says of itself that it is false. One of the earliest appearances of this sort of puzzle is in the Bible, where Paul of Tarsus relates an anecdote of a Cretan man who says that all Cretans are liars. This, by itself, is not so difficult to interpret in a way that dissolves the sense of paradox—perhaps this Cretan meant that all *other* Cretans are liars, or that all Cretans utter falsehoods some, but not necessarily all, of the time. Even under the simplifying assumption that the Cretan intended to suggest that every single

Cretan, including himself, utter falsehoods all the time, there is not yet a paradox. The sentence, to be sure, cannot be uttered truthfully. For if the sentence is true, then the Cretan must be lying, which means it must be false. But if the sentence is false, that just means that the speaker is a liar and there is some Cretan who is truth-telling, which is not a paradoxical conclusion.

But suppose instead of saying that all Cretans are liars, Paul's Cretan said it of himself or, more transparently, had simply said "this sentence is false" or had introduced a sentence *L* which said "*L* is false." In this case, it seems impossible to ascribe either truth or falsity to the sentence *L*. If *L* is true, then what it says must be the case, so it is false. If *L* is false, then it must not be the case that what it says is the case, so it is not false. Additionally, it seems to be no help to this paradoxical conclusion to allow for sentences to be neither true nor false. The sentence *L* can be modified to form the following sentence *L'*: "*L'* is either false or neither true nor false." There are three possibilities: *L'* is true, *L'* is false, or *L'* is neither true nor false. The first two possibilities lead to contradiction in exactly the same way that *L* did. But additionally, in the case that *L'* is neither true nor false, then what it says of itself is true, so it is true, hence can *not* be said to be neither true nor false. This leads once more to contradiction.

Russell's Paradox

Russell's paradox, introduced by the British philosopher and logician Bertrand Russell, is a paradox that concerns sets. Naively, a *set* is just a well-defined collection of objects. One way of describing a set is just to list its elements. For example, there's a set {1, 2, 3} which consists exactly of the elements 1, 2, and 3. But another way to describe a set is to give the conditions for membership in the set. For example, the set of physical objects that weigh less than a kilogram; it isn't possible to simply list all such objects, but there is a criterion for membership. To check whether some physical object is in the set or not, one can simply put it on a scale. In the context of mathematics, such ways of introducing sets—such as the set of prime numbers, the set of points on a given line, the set of deducible theorems from a given set of axioms—are completely ubiquitous.

Russell's paradox begins by considering the property of non-self-membership and considering the set of all sets with this property. Some sets are members of themselves and some are not. The set of all sets, assuming it exists, is a set, so it is a member of itself, while the empty set is a set with no members, so it is not a member of itself. So now consider the set *R* consisting exactly of those sets which are not members of themselves. We know already that the set of all sets is not in *R* and the empty set is in *R*. But is *R* a member of itself?

Herein lies the paradox: if *R* is a member of itself, then, by definition, it must have the property of non-self-membership and that means that *R* is not a member of itself. Therefore, *R* can't be a member of itself. But if *R* is *not* a member of itself, then it has the property of non-self-membership and consequently *R* must be a member of the set of all sets with the property of non-self-membership, which is to say *R* must be a member of *R*. We see, then, that if *R* is a member of itself, then it isn't, and if it isn't, then it is. This is a paradox.

Variations on Russell's paradox exist in natural language as well. For example, the well-known logic textbook *Model Theory* by the model theorists C. C. Chang and Jerome Keisler (2013) contains the following dedication: "For the amusement of all those who gave us help, we dedicate our book to all model theorists who have never dedicated a book to themselves" (p. xiv).

Did they dedicate the book to themselves? As in Russell's paradox, either way of answering this question seems to lead to contradiction. Another version was introduced by Kurt Grelling, who defined the adjective *heterological* to describe those adjectives that do not describe themselves. So, "monosyllabic" is heterological (because "monosyllabic" is not monosyllabic) but "polysyllabic" is not. Grelling's paradox arrives at the self-contradictory conclusion that *heterological* must be both heterological and not heterological.

Problems and Solutions

Willard van Orman Quine introduced a helpful way of distinguishing the varieties of paradoxes. If a paradox is a kind of argument that establishes an absurd conclusion, we can call this conclusion the *proposition* of the paradox. For example, the proposition of Zeno's paradox of Achilles and the

Tortoise is the proposition that Achilles never overtakes the tortoise. In other examples, the proposition of a paradox is somewhat less clear. For instance, Quine says of Grelling's paradox that its proposition is a self-contradictory compound proposition to the effect that our adjective is and is not true of itself. He then defines the *veridical* paradoxes to be the ones where the proposition is true, and the *falsidical* paradoxes to be the ones where the proposition is false, there being a fallacy at some place in the argument. Motion *is* possible, Achilles *does* in fact overtake the tortoise and fallacies can be found in each of the arguments. As a consequence, Zeno's paradoxes are falsidical. On the other hand, a well-known theorem of Stefan Banach and Alfred Tarski establishes the unit ball; that is, the area inside the sphere of radius 1 can be partitioned into finitely many pieces, moved around by rotations and translations in three-dimensional space, and thus rearranged to form two balls of the same volume as the first. This mathematical result is known as the *Banach–Tarski paradox*. The conclusion, despite flying in the face of our spatial intuition, is nonetheless a mathematical fact. This paradox, then, is veridical.

Quine's distinction classifies paradoxes *retrospectively*, from a point of view reached after the paradox has been solved. The veridical paradoxes are the ones where we resolve them by learning somehow to accept the intolerable conclusion. In the case of the Banach–Tarski paradox, for example, the absurdity of the conclusion is mellowed once one learns that the pieces of the partition are not obtained constructably and cannot be assigned volume—the components of the partition bear no resemblance to the kinds of pieces that could be physically realized by a decomposition of a ball in the world. The falsidical paradoxes are the ones we resolve by finding a fallacy in the argument that blocks the inference to the paradox's proposition. In the case of Achilles and the Tortoise, we noted that it doesn't follow that a duration of time which can be broken up into infinitely many periods of positive duration must therefore be infinite. Therefore, the inference that Achilles cannot reach the tortoise was based on bad reasoning and, once we give up this step in the argument, the paradox dissolves. But the paradoxes that are not yet resolved, for Quine, belong to a third category, the *antinomies*.

But how can a paradox be solved? One of the steps in Zeno's paradox, for example, was inferring from the fact that infinitely many tasks are performed, that infinitely much time must have passed. But this doesn't follow. If Task 1 takes $t/2$ many seconds, and Task 2 takes $t/4$ many seconds, and Task 3 takes $t/8$ many seconds, and so forth, then the amount of time taken up by performing Task n for all $n = 1, 2, \dots$ can be expressed by

$$\sum_{n=1}^{\infty} \left(\frac{t}{2}\right)^n = 1$$

using the sum-formula for a geometric series. So doing all the infinitely many tasks takes only t many seconds, which is a finite duration of time. This basic calculus fact intervenes to show us that the inference involved in Zeno's paradox was incorrect, and the paradox dissolves.

The solution of the Sorites paradox is considerably less straightforward. For this paradox, there is no consensus on even what constitutes the *source* of the paradox. In the paradox of the heap, it is fairly uncontroversial that a single grain does not constitute a heap, so treatments of the paradox tend to focus on the principal of tolerance being a heap is supposed to satisfy—namely, that a heap minus a single grain is still a heap. Some commentators believe that the concepts we employ when talking about heaps or other vague notions are indeterminate or incomplete. This requires, then, that some heap minus a single grain of sand becomes *indeterminate*. In order to avoid simply reproducing the paradox, then, these theories have to account for *higher order vagueness*, arguing that the concept of *determinate heap* is vague in the same way as *heap* is. Other theories argue that there, in fact, is a sharp boundary separating the heaps from the nonheaps, and the phenomenon of vagueness can be understood as a consequence of our *ignorance* of that boundary. On either account, the principle of tolerance associated to vague concepts is rejected or tweaked to block the inference that leads to contradiction in the Sorites paradox.

Why Paradoxes Matter for Science

Paradoxes, on first blush, appear to be more like riddles or sources of amusement rather than the objects of serious scientific inquiry. For example,

the Sorites paradox largely concerns concepts that occur in natural language, like “a heap of sand,” rather than more precisely quantifiable notions that appear in empirical science, like “greater than 50 grams of sand.” When vagueness would present a problem for the development of a scientific theory, reference to vague notions can often be replaced by vocabulary to which soritical difficulties do not apply. However, the notions of *truth* and *set*, called into question by the Liar paradox and Russell's paradox respectively, play a fundamental role in logic and the foundations of mathematics and cannot be so easily replaced or dispensed with. The urgency of dealing with these fundamental paradoxes in formal theories of mathematics, logic, and language has motivated several major conceptual advances in each of these realms.

Russell's paradox was discovered by Bertrand Russell, who came across the paradox while trying to make sense of Gottlob Frege's work in the logicist program of reducing mathematics to logic. Frege had worked out an elaborate formal development of mathematics based on logical principles, but one of the axioms of his system was his Basic Law V, which roughly asserts, in contemporary language, that, for every property, there is a set consisting exactly of the objects that possess that property. For example, the property of being a prime number can be used to form the set of prime numbers by application of this principle. Russell observed that paradox results by applying this principal to the property of non-self-membership as described above. Upon learning of the paradox, Frege felt that a contradiction had been discovered at the heart of mathematics itself: His reply to Russell (quoted in Quine 1966) said “arithmetic totters.”

In order to avoid this paradox, Russell and Alfred North Whitehead, in their *Principia Mathematica*, introduced a theory of *types*. In simplified form, the objects of Type 0 are the individuals, objects of Type 1 are collections of individuals, objects of Type 2 are collections of collections of individuals (or, equivalently, collections of objects of Type 1), and so on. In this system, an object of Type n could only be an element of an object of Type $n + 1$, which means that no object could ever be an element of itself. In this way, the paradox was blocked, although at the cost of working

with a cumbersome formalism. Later approaches to the foundations of mathematics, by way of Zermelo-Frankel set theory, eschew the theory of types but prohibit sets being members of themselves. In particular, in these systems, there can be no set of all sets. Finding a system that could serve as the foundations of mathematics was a central research program within 20th-century mathematics, and Russell's paradox was a central consideration in the way the formalism eventually took shape.

The Liar paradox, likewise, has had a substantial influence on the way truth has been understood within logic and mathematics. Avoiding antinomies like the liar motivated Alfred Tarski to introduce a definition of truth in formal languages, which has since become a fundamental notion in model theory and plays an important role in linguistics in the area of formal semantics. A key feature of Tarski's approach was that truth for expressions in a language L being analyzed, called the *object* language, while the definition itself was expressed in a different, possibly richer language called the *metalinguage*. In this approach, expressions about the truth of sentences in L cannot be expressed in L itself, so the liar sentence cannot be formulated. Indeed, by making use of a formalized version of the Liar paradox, Tarski showed that no sufficiently rich formal language can contain its own truth predicate; he showed that, within the object language, it is not possible to define the set of true sentences in that language. An alternative approach, in which the object language may contain its own truth predicate but truth and falsity are only applied to *some* of the sentences, was pursued by Saul A. Kripke and the resulting theory was very influential for both philosophy and mathematical logic.

The aforementioned advances were motivated by a desire to create formal systems that could capture the features of truth and our reasoning about sets while eliminating the paradoxes, but one further application—perhaps the most striking of all of them—involved finding the liar paradox where it was not expected. Gödel's incompleteness theorem, originated by Kurt Gödel, addresses the question of whether one could discover all of the truths of arithmetic. For example, children learn in school that “for every

number n and every number m , $n + m = m + n$ ” and, perhaps later, learn how to apply the principle of induction to prove formulas like

$$1 + 2 + \dots + n = n(n + 1)/2$$

which holds for all natural numbers n . Clearly, there are infinitely many such truths, but it is natural to ask if there are finitely many principles from which all such truths could be deduced or, perhaps, if one could program a computer to begin churning out a list of truths about the natural numbers so that, given infinitely much time, every truth would eventually appear. In contemporary language, Gödel showed how any collection of arithmetic truths that could be described by a computer program must be incomplete. And he did this by showing how, from such a set of truths, one could produce a statement that said of itself “I am not provable from this set of statements.” This is exactly analogous to the liar sentence, except with provability from the axioms taking the role of truth. If Gödel's statement were provable, then, assuming the set of statements were consistent, it would be true, hence not provable. This shows this sentence is *not* provable and thus that the set of provable statements must be incomplete. In other words, there must be some arithmetical truths that elude our grasp, even given arbitrary computational strength and time. In this way, Russell's paradox and the Liar paradox have underlain some of the most significant developments in philosophy, logic, and mathematics.

Nicholas Ramsey

See also Induction; Logic and Language; Logic, Formal and Informal; Logical Theory; Modeling; Set Theory

Further Readings

- Banach, S., & Tarski, A. (1924). Sur la décomposition des ensembles de points en parties respectivement congruentes [On the decomposition of sets of points into respectively congruent parts]. *Fundamenta Mathematicae*, 6(1), 244–277.
- Chang, C. C., & Keisler, H. J. (2013). *Model theory* (3rd ed.). Dover Publications.
- Feferman, S. (1991). Reflecting on incompleteness. *The Journal of Symbolic Logic*, 56(1), 1–49.

- Feferman, S. (2008). Axioms for determinateness and truth. *The Review of Symbolic Logic* 1(2), 204–217.
- Friedman, H., & Sheard, M. (1987). An axiomatic approach to self-referential truth. *Annals of Pure and Applied Logic*, 33, 1–21.
- Hardie, R. P., & Gay, R. K. (1933). Aristotle's physics. *Philosophical Review*, 42(3). 322–323.
- Kripke, S. A. (1980). *Naming and necessity*. Harvard University Press.
- Nelson, L., & Grelling, K. (1908). Bemerkungen zu den paradoxien von Russell und Burali-Forti [Remarks on the paradoxes of Russell and Burali-Forti]. *Abhandlungen der Fries'schen Schule. Neue Folge*, 2, 300–334.
- Plutarch (n.d.). *Plutarch's lives (Volumes I and II)*. (J. Dryden, trans.). <https://gutenberg.org>
- Quine, W. V. (1966). *The ways of paradox*. Harvard University Press.
- Tarski, A. (1936). The concept of truth in formalized languages. In Tarski, A. *Logic, semantics, metamathematics* (John Corcoran, ed., J. H. Woodger, trans.) (pp. 152–278). Oxford University Press.
- van Heijenoort, J. (2002). *From Frege to Gödel: A source book in mathematical logic* (pp. 1879–1931). Harvard University Press.
- Whitehead, A. N., & Russell, B. (1927). *Principia mathematica*. Cambridge University Press.

PERTURBATION THEORY

This entry aims to explain essential elements of perturbation theory and their conceptual underpinnings. It is not meant as a summary of popular perturbation methods, though some illustrative examples are given to underline the main methodological insights and concerns. We also give brief explications of the mathematical notions of *limit*, *continuity*, *differentiability*, *convergence*, and *divergence*, which provide the necessary foundation.

The role of perturbation theory is to study the effects of small disturbances (or perturbations) on the behavior of a (possibly only hypothetical) system irrespective of the source of the disturbance. For instance, the disturbance may have its physical origin in the environment of a system, as when a system receives small amounts of energy from the outside or when an experimental setup is slightly disturbed by the vibration caused by a

truck on a nearby street. The disturbance may have been added to the modeling equations by design in order to simplify the study of a system, as is the case when we neglect small quantities to obtain more mathematically tractable equations. The disturbance may even be due to experimental error and uncertainty or to computational error.

The effects of small disturbances are sometimes small too, in which case they are usually called *regular* perturbations. As James G. Simmonds and James E. Mann (1998) explain, regular perturbations “are assumed nearly every time we construct a mathematical model of a real world phenomenon. Our choice of language reflects this: the flow is *almost* steady, the density varies *essentially* with altitude only, the conductivity is *virtually* independent of temperature, the spring is *nearly* linear, the friction is *practically* negligible” (p. 1). However, mother nature isn’t always that cooperative and sometimes the effect of small disturbances is very large indeed; in such cases, we talk of *singular* perturbations. We will return to the distinction between regular and singular perturbations later. Perturbation theory is central to the scientific method in two important but distinct respects. On the one hand, good applied mathematical practice demands that one considers possible disturbances before assessing whether a mathematical model correctly captures the behavior of a system. Thus, even “exact” mathematical theories and models enter the realm of perturbation theory as soon as they are applied. On the other hand, while it is typical of introductory courses on “mathematical X”—for example, mathematical physics, biology, economics—to focus exclusively on problems that can be solved exactly, this desirable event of complete tractability only rarely happens in applied mathematical practice, either because we don’t know exactly solvable models or because the accurate models turn out to be theoretically impossible to solve exactly. Hence instead of using “demonstrative solution methods” one necessarily relies on methods that extract approximate solutions. In this case, a problem is said to be “solved” in the sense that we can get approximations that are close enough, even if we can never actually solve it exactly (we return to this below).

One of the beautiful and powerful aspects of perturbation theory is that it addresses the two concerns with the very same set of mathematical methods. These methods, to be sure, differ from

other mathematical methods that are central to science. For instance, deductive or axiomatic approaches focus on whether axioms are true and on whether certain statements are logical consequences of sets of axioms. Probability theory concentrates on quantifying how probable it is that a statement will be true. However, the few remarks above suggest that in most concrete contexts, it is not enough to ask what is true, what logically follows from what, and what is probable; we also need to ask: *If things were changed or disturbed a little, what would the effect be?* This is precisely the role of perturbation theory and what distinguishes it from other mathematical disciplines.

Mathematical Background

To understand the main concepts and methods of perturbation theory, we introduce a few essential mathematical ideas. We believe that they will be accessible to most readers who have an intuitive grasp of the concept of derivative and, to a lesser extent, of the concept of integral.

We begin with the crucial notion of *limit* and the related notion of *convergence*. Intuitively, a limit is the mathematical object (often, a number or a function) that is approached by an infinite sequence. Consider a classical example. Even if we know by definition that π is the ratio of the circumference of a circle to its diameter, it is impossible to write down its value as a number in conventional format since its decimal expansion is infinite and nonrepeating. However, around 250 BCE, Archimedes made the first known use of sequences to calculate the value of π to arbitrary precision, at least in principle.

As we see in Figure 1, the perimeter of the hexagon circumscribed in the unit circle is equal

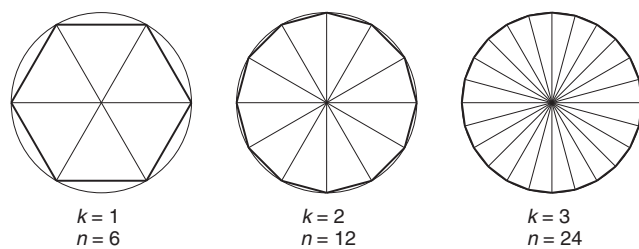


Figure 1 The First Three Items in Archimedes' Sequence of Polygons Circumscribed in the Unit Circle to Approximate π

to 6, since it can be decomposed in six equal equilateral triangles of Side 1. Archimedes' brilliant idea was that, if we know the side s of a polygon circumscribed in a circle, we can find the side s' of a polygon with twice the number of sides by elementary trigonometry (see Figure 2):

$$s' = \sqrt{\left(1 - \sqrt{1 - \left(\frac{s}{2}\right)^2}\right)^2 + \left(\frac{s}{2}\right)^2} \quad (1)$$

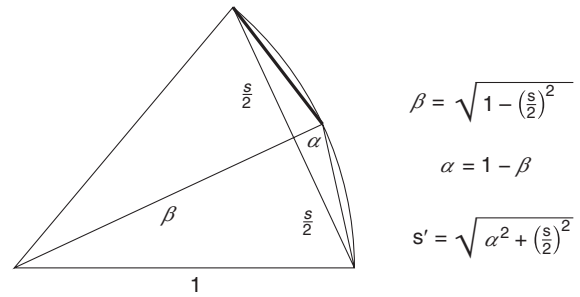


Figure 2 The Trigonometric Relation Exploited by Archimedes

Using this formula, we can successively find the length of the sides of circumscribed polygons with 6, 12, 24, 48, . . . sides, from which it is easy to find approximations of the ratio of the perimeter to the diameter. This gives us the following sequence of approximations:

$$\begin{aligned} p_1 &= 3.0000 & p_2 &= 3.1058 & p_3 &= 3.1326 & p_4 &= 3.1394 \\ p_5 &= 3.1410 & p_6 &= 3.1415 & p_7 &= 3.1416 & & \dots \end{aligned}$$

Here, we round to the fourth decimal place for convenience. The bigger k is, the closer p_k will be to π . In that sense, we say that the sequence of p_k converges to π as k goes to infinity (denoted $k \rightarrow \infty$). Alternatively, we say that π is the limit of this sequence as $k \rightarrow \infty$.

Limits of sequences are also essential to the manipulation and interpretation of *infinite series*:

$$S = \sum_{k=0}^{\infty} a_k = a_0 + a_1 + a_2 + \cdots + a_n + \cdots \quad (2)$$

Consider the perfectly well-defined finite sum

$$S_N = \sum_{k=0}^N a_k = a_0 + a_1 + a_2 + \cdots + a_N \quad (3)$$

of $N + 1$ terms. The sums $S_1, S_2, S_3, \dots$ form a sequence. The sum S of an infinite series is interpreted as the limit of this sequence, that is, the number or function that is approached as the number $N + 1$ of terms summed approaches infinity. In mathematical notation, we write

$$S = \lim_{N \rightarrow \infty} S_N = \lim_{N \rightarrow \infty} \sum_{k=0}^N a_k = \sum_{k=0}^{\infty} a_k. \quad (4)$$

If no such finite S exists, the series is said to *diverge*. However, if the series has a limit S , then it is possible to use a finite number N of terms to approximate the value of S . Since this practice involves retaining the first N terms of the series and discarding the rest, this is called a *truncated series approximation*.

There is a substantive arsenal of tests available to determine whether series converge. But what is the status of series that don't converge to any number or function? In the 19th century, Niels Henrik Abel provocatively put it as follows:

The divergent series are the invention of the devil, and it is a shame to base on them any demonstration whatsoever. By using them, one may draw any conclusion he pleases and that is why these series have produced so many fallacies and paradoxes [...]. (cited in Hoffman, 1998, p. 218)

From the point of view of analysis, they are said to have no meaning since they don't refer to anything; alternatively, the series is said not to be valid.

However, as we will see, there are important differences between this perspective and that which provides the conceptual foundations for perturbation theory. Indeed, additional subtleties come into play when we consider *series of functions* that may (or may not) converge to a function, and many of the most successful applications of mathematics in science have only become possible by an improved understanding of how divergent series can be used to obtain reliable results.

First, we observe that functions themselves may have a limit near a point; this is an application of the same idea we've introduced above. Consider a function $f(x)$ defined on an interval (or more generally, on an open set) A , and suppose we wish to know what value the function $f(x)$ approaches as

the argument x becomes increasingly closer to a point a within the interval. We can think of a sequence of arguments $a_0, a_1, a_2, a_3, \dots$ and of the corresponding sequence of values $f_1 = f(a_1), f_2 = f(a_2), f_3 = f(a_3), \dots$. If the sequence of k s approaches a given finite value L as the k s approach a , for any choice of sequence of k s that approach a , then we say that L is the limit of the function $f(x)$ as x approaches a . In rigorous mathematical notation, we say that a function $f(x)$ has the limit L as $x \rightarrow a$ if

$$(\forall \epsilon > 0)(\exists \delta > 0)(x \in A \ \& \ |x - a| < \delta \Rightarrow |f(x) - L| < \epsilon). \quad (5)$$

A function $f(x)$ is said to be *continuous* at the point $x = a$ provided that

$$\lim_{x \rightarrow a} f(x) = f(a). \quad (6)$$

It is continuous on an interval (or open set) A if it is continuous at all points a in the interval. This definition captures the intuitive idea that a function is continuous if its graph is an unbroken curve, without holes or jumps.

The treatment of limits at ∞ is nearly the same, because the ideas are very similar (some would say identical except in form). We say that

$$\lim_{x \rightarrow \infty} f(x) = L \quad (7)$$

if

$$(\forall \epsilon > 0)(\exists X)(x > X \Rightarrow |f(x) - L| < \epsilon). \quad (8)$$

Here, the large X plays the same role as the small δ earlier.

This is known as the *epsilon-delta* definition of limit and it is due to the efforts of Cauchy, Bolzano, and Weierstrass to provide rigorous foundations for calculus in the early 19th century. This definitional approach provides concrete meanings for more intuitive phrases, used since the time of Newton, such as “we can make the error $|f(x) - L|$ as small as we please by taking x sufficiently close to a ,” without appealing to infinities and infinitesimals.

Following Carathéodory, we say that f is *differentiable* at $x = a$ if there exists a function $\phi(x)$ that is continuous at $x = a$ such that

$$f(x) = f(a) + (x - a)\phi(x), \quad (9)$$

in which case $f'(a) = \lim_{x \rightarrow a} f(x) = f(a)$. This is of course the equation of the tangent of the function f at the point $x = a$. The process can be repeated to find the derivative of the derivative (known as *the second derivative*) $f''(x)$, the third derivative $f'''(x)$, and so on. This gives rise to an *expansion* of the function f about the point a . When a function has $N + 1$ derivatives in a neighborhood of $x = a$, then there exists a continuous function $\phi_{N+1}(x)$ such that

$$= \sum_{k=0}^N \frac{1}{k!} f^{(k)}(a)(x-a)^k + \phi_{N+1}(x)(x-a)^{N+1}. \quad (10)$$

This is known as *Taylor's theorem*, and the resulting Taylor series are among the most useful in applied mathematics.

To simplify expressions such as the one in Equation 10, we can use the “big-Oh” notation, first used by Bachmann and popularized by Landau. To indicate that $f(x)$ and $g(x)$ are of similar size near $x = a$, we say that f is order g (denoted $f(x) = O(g(x))$ near a). More precisely, we say that

$$f(x) = O(g(x)) \quad \text{near } x = a \quad (11)$$

if there exists a constant K such that for all x sufficiently near a we have

$$|f(x)| \leq K|g(x)|. \quad (12)$$

In theory K may be very large—indeed arbitrarily large or arbitrarily small. But in practice, when the use of a perturbation method is successful, K is usually of quite moderate size. With this notation, Equation 10 simply becomes

$$f(x) = \sum_{k=0}^N \frac{1}{k!} f^{(k)}(a)(x-a)^k + O((x-a)^{N+1}). \quad (13)$$

Following the same idea, but generalizing it, perturbation theory uses truncated series of functions:

$$S_N(\varepsilon) = \sum_{n=0}^N a_n \varphi_n(\varepsilon). \quad (14)$$

The sequence $\varphi_0(\varepsilon), \varphi_1(\varepsilon), \dots, \varphi_N(\varepsilon), \varphi_{N+1}(\varepsilon), \dots$ may be infinite; these functions are often called *gauge functions*. The most usual gauge functions are simple powers of $1, \varepsilon, \varepsilon^2, \dots$ (i.e., ...) when we are interested by the behavior of the function as $\varepsilon \rightarrow 0$, or powers of $(\varepsilon - a)$ when we are interested

in the behavior of the function as $\varepsilon \rightarrow a$. Notice that when ε is near 0 or a , as the case may be, the higher the power N of ε^N or $(\varepsilon - a)^N$ is, the closer it will be to 0; in other words, for gauge functions, each function in the sequence is smaller than the previous one near the limit point:

$$\lim_{\varepsilon > 0} \frac{\varphi_{n+1}(\varepsilon)}{\varphi_n(\varepsilon)} = 0 \quad (15)$$

This is only useful if the φ_k are somehow simpler than the original function f . Indeed we know an example,

$$\ln z = \sum_{k=1}^N W^{(k)}(z) + O(W^{(N+1)}(z)) \text{ as } z \rightarrow 0 \quad (16)$$

of an asymptotic expansion of a simple function in terms of iterates of a less well-known function, the Lambert W function, but if this is useful at all the example would be a startling exception (see Corless et al., 1996). The normal case has each φ_k simpler than the original f .

One might also say that perturbation methods are an application of the theory of truncated *formal* series. If one does not explicitly truncate the series as in Equation 14 but writes $a_0\varphi_0(\varepsilon) + a_1\varphi_1(\varepsilon) + a_2\varphi_2(\varepsilon) + \dots$, one describes the series as “formal.” This means that issues of convergence as $N \rightarrow \infty$ are usually irrelevant, and in practice, divergent series are often used. The point is, and it is a simple one although there is a great deal of confusion in the literature, that there are two limits of interests with series:

$$\lim_{N \rightarrow \infty} S_N(\varepsilon), \text{ where here } \varepsilon \text{ is fixed, and } \varepsilon \neq 0,$$

encountered first in most calculus classes, whence issues of convergence of a sum of an infinite number of terms arise, and the other

$$\lim_{\varepsilon \rightarrow 0} S_N(\varepsilon), \text{ where here } N \text{ is fixed,}$$

the relevant limit for perturbation series. In the quite common case that the second limit as $\varepsilon \rightarrow 0$ exists and is finite but the first does not (the series diverges as $N \rightarrow \infty$) the series is said to be “asymptotic.” The cure for divergence as $N \rightarrow \infty$ is simply

to keep N fixed. The other limit, as $\varepsilon \rightarrow 0$, often provides all the accuracy one needs.

This point of view allows one to use a fixed number of terms (*any* fixed number of terms) of a series that would diverge if we were so foolish as to let $N \rightarrow \infty$. Using gauge functions, that is, functions that are smaller than the previous in a neighborhood of the limit point, we find that perturbation series for $f(a + \varepsilon)$ is

$$f(a + \varepsilon) = \sum_{k=0}^N p_k \varphi_k(\varepsilon) + O(\varphi_{N+1}(\varepsilon)) \quad (17)$$

if

$$\lim_{\varepsilon \rightarrow 0} \frac{|f(a + \varepsilon) - \sum_{k=0}^m p_k \varphi_k(\varepsilon)|}{|\varphi_{m+1}(\varepsilon)|} \quad (18)$$

is bounded (*mutatis mutandis* as $t \rightarrow \infty$), for each $m = 0, 1, 2, \dots, N$. Notice that N is fixed and does not go to infinity.

The Use of Series in Perturbation Theory

With the mathematical concepts introduced in the previous section, we are now ready to examine how perturbation methods work. When facing a hard problem, one can use the following conceptually (and with some practice, mathematically) simple three-step method. The *first step* consists in introducing a small positive parameter ε , known as a *perturbation parameter*, in the hard problem that you may not know how to solve or that can be provably unsolvable by explicit methods. So, instead of having only one hard problem, we are now in the presence of an infinite collection of hard problems, one for each value of ε . This may not look like progress, but that is only the first step. The idea of introducing a perturbation ε is sometimes an obvious thing to do (e.g., when a value of ε such as 0 really simplifies things), and sometimes it is not obvious at all. Very clever choices are sometimes needed, and knowing what choice to make is in many respects an art rather than a science. However, in scientific use, the modeling context may sometimes dictate where to introduce the ε as it will correspond to a quantity that is small, such as air resistance, friction, and so on. For the *second step*, assume that the answer is a function of ε and assume that

the answer is in the form of a truncated power series in ε :

$$ANS(\varepsilon) = \sum_{n=0}^N \alpha_n \varepsilon^n \quad (19)$$

More generally, we could use any collection of gauge functions instead of the powers of ε . This practice, which seems to assume that we have a valid series expansion without actual mathematical evidence, is sometimes objected to; but we will explain below the sense in which it is legitimate. The main problem consists in calculating the coefficients α_n , and our examples below show how this can be achieved recursively. This is why perturbation theory is sometimes defined as follows:

Perturbation theory is a large collection of iterative methods for obtaining approximate solutions to problems involving a small parameter ε . (Bender & Orszag, 1978, p. 319)

Note also that since we aim to obtain a truncated series approximation, we will fix the number of coefficients that we will solve for (N will often be only 1 or 2, in which case we will say, respectively, that we obtained a first-order or second-order solution).

In the third and final step, we set ε to the value that recovers the original problem we wanted to solve (often, it will be $\varepsilon = 1$) and just calculate the approximate solution by summing a finite number of terms of the perturbation series. That simple method very often works, but as we will see, sometimes more care is needed. We begin with an example in which it turns out to work well.

Consider the equation $x^5 + x - 1 = 0$. Our problem is to find a real root of this equation, that is, a value of x that satisfies this equation. The problem is hard, but we need not despair as we can use the method outlined above. Firstly, we introduce an ε as follows:

$$x^5 + \varepsilon x - 1 = 0 \quad (20)$$

We will choose to keep only two terms to obtain a first-order approximation. We thus assume that our truncated series solution will have the form $x(\varepsilon) = a_0 + a_1 \varepsilon$. We substitute this expression in Equation 20:

$$(a_0 + a_1 \varepsilon)^5 + \varepsilon(a_0 + a_1 \varepsilon) - 1 = 0 \quad (21)$$

We expand the fifth-power term using Pascal's triangle:

$$(a + b)^5 = a^5 + 5a^4b + 10a^3b^2 + 10a^2b^3 + 5ab^4 + b^5 \quad (22)$$

After gathering like terms, we obtain

$$a_0^5 - 1 + (5a_0^4 a_1 + a_0) \varepsilon + (10a_0^3 a_1^2 + a_1) \varepsilon^2 + O(\varepsilon^3) = 0. \quad (23)$$

hiding the complicated higher-order terms in the O -symbol (we do not need them anyway, since the ε^2 term will already fail to be 0 with this first-order approximation). We then match the coefficients of matching powers of ε on the left-hand side and the right-hand side. Since all coefficients of ε on the right-hand side are 0, we obtain the following equations for the first two coefficients:

$$a_0^5 - 1 = 0 \quad (24)$$

$$5a_0^4 a_1 + a_0 = 0 \quad (25)$$

Since we were interested with the real root of this equation, we obtain the value $a_0 = 1$ from Equation 24 and $a_1 = -1/5$ from Equation 25. We see that the solution of the coefficients is a recursive process that is in itself a purely mechanical task that presents no essential mathematical difficulty, provided the zeroth order equation, which is generally nonlinear, can be solved.

Thus, our approximate solution, which we denote $z(\varepsilon)$ in contrast with the exact but unknown solution $x(\varepsilon)$, can be written as follows:

$$z(\varepsilon) = 1 - \frac{\varepsilon}{5} \quad (26)$$

Notice that in Equation 20, if we set $\varepsilon = 1$, we obtain our original unperturbed problem. Thus, our approximate root will be $z(1) = 1 - 1/5 = 0.8$; it does not compare too badly with the exact solution, which is 0.755 rounded to three digits. If we were to do a bit more work, we would find that 2; accordingly our second-order approximation would be 0.76, even better!

Notice that we have provided no evidence to justify any step in which we assumed that the root $x(\varepsilon)$ of Equation 20 can be meaningfully written as an infinite power series. But in fact, we have

made no such assumption! The approach here can be justified by an explicitly finite approach borrowed from numerical analysis that in many cases allows physical interpretation, based on the concept of *residual* (see, e.g., Corless & Fillion, 2013).

The idea is quite general; the unknown function $f(\varepsilon)$ usually satisfies a known equation $\Phi(f; \varepsilon) = 0$ by definition (above, we had $\Phi(f; \varepsilon) = \Phi(x, \varepsilon) = x^5 + \varepsilon x - 1 = 0$). However, if we evaluate Φ using z instead of x , we wouldn't obtain 0 since z isn't the exact solution. The residual Δ is the quantity we obtain once we substitute the exact function $f(\varepsilon)$ by its truncated series approximation:

$$\Delta := \Phi \sum_{k=0}^N a_k \varepsilon^k, \varepsilon \quad (27)$$

Then z satisfies

$$\Phi(z; \varepsilon) - \Delta = 0,$$

not $\Phi(z; \varepsilon) = 0$.

In general if we correctly identify the first $N + 1$ coefficients, we obtain a residual that satisfies

$$|\Delta| = O(\varepsilon^{N+1}) \quad (28)$$

as $\varepsilon \rightarrow 0$. As we have explained, the symbol $O(\varepsilon^N)$ is conveniently used to refer to a function that decays like ε^N without having to specify it explicitly. In the example above, we implicitly used the notion of residual to find coefficients that guarantee that Equation 28 will be satisfied, that is, that the residual will be of higher order than the terms that are part of our approximate solution. Instead of writing Equation 23, we could more exactly have written our equation for the residual,

$$\Delta = a_0^5 - 1 + (5a_0^4 a_1 + a_0) \varepsilon + (10a_0^3 a_1^2 + a_1) \varepsilon^2 + O(\varepsilon^3), \quad (29)$$

and by an iterative inspection of the formula, we find that choosing $a_0 = 1$ guarantees that the residual is $O(\varepsilon^2)$ irrespective of the other coefficients, and that in addition choosing $a_1 = -1/5$ guarantees that the residual is. For small ε , this z is the exact solution of nearly the right equation. By doing more work, but using no new ideas, we may find $z_N = a_0 + a_1 \varepsilon + a_2 \varepsilon^2 + \dots + a_N \varepsilon^N$ for any fixed $N \geq 1$ with $\Delta_N = z_N^5 + \varepsilon z_N - 1 = O(\varepsilon^{N+1})$,

which can be made arbitrarily small by taking ε arbitrarily close to 0. Also, for this example, more is true. We may in this case take $N \rightarrow \infty$, and find an explicit formula for all a_k , thereby giving an infinite series—which turns out to be convergent—and the exact root. But that is no longer perturbation theory.

The question of which number N of terms is best can be answered easily: N is best when the magnitude of the residual $|\Delta|$ is smallest. Then, the error $|z - x|$ satisfies the inequality

$$|z - x| \leq \kappa |\Delta| \quad (30)$$

and $|z - x|$ will also be smallest. Here, κ is the condition number of the problem (see, e.g., Corless & Fillion, 2013), a quantity that is essentially a bound on

$$\left| \left(\frac{\partial \Phi}{\partial x} \right)^{-1} \right| \quad (31)$$

Singular Perturbations

A regular perturbation problem may be regarded as one for which the character of the solution set to $\Phi(f, 0) = 0$ is the same as the character of the solutions to $\Phi(f, \varepsilon) = 0$ for small ε . A singular perturbation problem is one for which the character changes discontinuously at $\varepsilon = 0$ (*mutatis mutandis*, $t = \infty$). The position of the perturbation parameter may play a key role in determining whether a perturbation is regular or singular, as these examples show:

Regular	Singular
$x^2 - 1 - \varepsilon = 0$	$\varepsilon x^2 + x - 1 = 0$
$y'' + \varepsilon y' + y = 0$	$\varepsilon y'' + y' + y = 0$
$y(0) = 1, y(1) = -1$	$y(0) = 1, y(1) = -1$

There are various definitions of singular perturbation problems in the literature, such as the one by Carl Bender and Steven Orszag (1978): “We define a singular perturbation problem as one whose perturbation series either does not take the form of a power series, or, if it does, the power series has a vanishing radius of convergence” (p. 324). They also use the more suggestive language: “[In either case], the exact solution for $\varepsilon = 0$ is *fundamentally different in character* from

the ‘neighbouring’ solutions obtained in the limit as $\varepsilon \rightarrow 0$.”

But this slippery distinction may not always be critical for understanding. Indeed, one common technique for solving so-called singular perturbation problems is to transform them (e.g., by change of variable) into regular perturbation problems. Interestingly, following this characterization, problems with spurious secular terms in the regular expansion display some aspects of both regular and singular perturbation problems (more on this in the example of Equation 45).

As a first singular example, consider $\Phi(x, \varepsilon) = \varepsilon^4 x^5 + x - 1 = 0$. This is a prototypical singular perturbation problem. For $\varepsilon \neq 0$, there are five (complex) roots. However, for $\varepsilon = 0$, there is only one, namely $x_0 = 1$. The character—the number of roots—has changed.

We may find the root near $x = 1$ by the same method as the previous example: let $z(\varepsilon) = a_0 + a_1 \varepsilon + a_2 \varepsilon^2 + a_3 \varepsilon^3 + a_4 \varepsilon^4 + a_5 \varepsilon^5$ and set as many coefficients of powers of ε in $\Delta = \varepsilon^4 z^5 + z - 1$ to zero as possible. This gives $z = 1 - \varepsilon^4 + \dots$ (implicitly stopping at some finite N). The other roots are more interesting. To find perturbation series for them, we transform to a regular problem, by using the change of variable $y = \varepsilon x$. Multiplying the original equation by ε gives

$$\varepsilon^5 x^5 + \varepsilon x + \varepsilon = 0 \quad (32)$$

and doing the change of variable gives

$$y^5 + y - \varepsilon = 0 \quad (33)$$

This is a regular perturbation problem for y , with five roots. By the same method as before, perturbation series for each can be found:

$$y_k(\varepsilon) = e^{(2k+1)\pi i/4} + \frac{1}{4e^{(2k+1)\pi i} + 1} \varepsilon + O(\varepsilon^2) \quad (34)$$

for $0 \leq k \leq 3$ and $y_4(\varepsilon) = \varepsilon - \varepsilon^5 + O(\varepsilon^9)$, giving

$$x_k(\varepsilon) = \frac{e^{(2k+1)\pi i/4}}{\varepsilon} + \frac{1}{4e^{(2k+1)\pi i} + 1} O(\varepsilon) \quad (35)$$

and

$$x_4(\varepsilon) = 1 - \varepsilon^4 + O(\varepsilon^8) \quad (36)$$

as perturbation series solutions for the original (singular) perturbation problem. Notice in this case that four of the solutions go to (complex) infinity as $\varepsilon \rightarrow 0$, leaving only the root near 1

being finite. This observation gives us some geometrical understanding of the reason for which this perturbation problem is singular.

Now, let us examine a problem that turns out to be well-behaved even though we may superficially expect a singular behavior. Consider

$$I(\varepsilon) = \int_0^{\infty} \frac{e^{-x}}{1 + \varepsilon x} dx \quad (37)$$

This integral exists for $\varepsilon \geq 0$. $I(\varepsilon)$ can be expressed in terms of special functions (so-called exponential integrals) but a simpler expression of its behavior for small ε would be helpful. Naive manipulations give

$$\frac{1}{1 + \varepsilon x} = 1 - \varepsilon x + (\varepsilon x)^2 + \dots \quad (38)$$

by the geometric series, so that

$$I(\varepsilon) = \int_0^{\infty} (1 - \varepsilon x + \varepsilon^2 x^2 + \dots) e^{-x} dx, \quad (39)$$

suggesting that

$$\begin{aligned} I(\varepsilon) &\doteq \int_0^{\infty} e^{-x} dx - \varepsilon \int_0^{\infty} x e^{-x} dx + \varepsilon^2 \int_0^{\infty} x^2 e^{-x} dx + \dots \\ &= 1 - 1!\varepsilon + 2!\varepsilon^2 + \dots \end{aligned} \quad (40)$$

It may seem remarkable that these naive manipulations, walking over the precipice of divergence for $|x| > 1$ as they do, manage to give useful perturbative expansions for $I(\varepsilon)$. Let us do these manipulations more carefully to see why. Note that

$$\frac{1}{1 + \varepsilon x} = \frac{1 - (-\varepsilon x)^{N+1}}{1 - (-\varepsilon x)} + \frac{(-\varepsilon x)^{N+1}}{1 + \varepsilon x} \quad (41)$$

for all x such that we avoid the singularity as $\varepsilon x = -1$; if we restrict $\varepsilon > 0$, then $1 + \varepsilon x \geq 1$ for $x \geq 0$ so division by zero never occurs. Now,

$$\frac{1 - (-\varepsilon x)^{N+1}}{1 - (-\varepsilon x)} = 1 - \varepsilon x + \varepsilon^2 x^2 - \dots + (-\varepsilon x)^N \quad (42)$$

is the sum of the *finite* geometric series, so

$$I(\varepsilon) = \int_0^{\infty} \left(\sum_{k=0}^N (-1)^k \varepsilon^k x^k \right) e^{-x} dx + \int_0^{\infty} \frac{(-\varepsilon x)^{N+1} e^{-x}}{1 + \varepsilon x} dx$$

$$\begin{aligned} &= \sum_{k=0}^N (-1)^k \varepsilon^k \int_0^{\infty} x^k e^{-x} dx + \int_0^{\infty} (-\varepsilon)^{N+1} \frac{x^{N+1} e^{-x}}{1 + \varepsilon x} dx \\ &= \sum_{k=0}^N (-1)^k \varepsilon^k k! + (-1)^{N+1} \varepsilon^{N+1} \int_0^{\infty} \frac{x^{N+1} e^{-x}}{1 + \varepsilon x} dx \end{aligned} \quad (43)$$

where all manipulations are of finite quantities and hence legitimate. Now, we do *not* let $N \rightarrow \infty$; rather, we look at $\varepsilon \rightarrow 0$. The term $\int_0^{\infty} \frac{x^{N+1} e^{-x}}{1 + \varepsilon x} dx$ is bounded (by $(N+1)!$) and hence the error term is $O(\varepsilon^{N+1})$ as $\varepsilon \rightarrow 0$, a term of higher order than all retained terms. Thus, this finite sum is a valid perturbation expression for $I(\varepsilon)$, in spite of the fact that if we were foolish enough to let $N \rightarrow \infty$, we would have a divergent series unless $\varepsilon = 0$, as can easily be seen for example by the ratio test.

As a practical matter, $I(0.01)$ can be evaluated quite accurately using only three terms of the series:

$$I(0.01) \doteq 1 - 0.01 + 2 \cdot 0.01^2 \doteq 0.9902 \quad (44)$$

The exact value is $I(0.01) = 0.9901942287$, so we see that we have obtained a very accurate answer with very little work. This exemplifies why convergence as $N \rightarrow \infty$ is of little or no importance in perturbation theory; what matters is that the finite perturbation series be continuous at $\varepsilon = 0$, indeed differentiable. Of course, this series is useless for large values of ε .

Finally, we discuss an example that illustrates the fact that problems whose perturbative solutions contain secular terms share features of both regular and singular perturbation problems. The equation

$$y'' + (1 - \varepsilon x)y = 0 \quad (45)$$

with initial conditions $y(0) = 1$, $y'(0) = 0$ is an example of a problem that is a *regular* perturbation problem on intervals $0 \leq x \leq X$ where $X = O(\varepsilon^{-\frac{1}{2}})$, that is, such that X is smaller than constant times $1/\sqrt{\varepsilon}$ when ε is small. Indeed, a regular perturbation expansion gives

$$\begin{aligned} y(x) &= \cos(x) + \varepsilon \left(\frac{1}{4} x \cos(x) + \left(-\frac{1}{4} + \frac{1}{4} x^2 \right) \sin(x) \right) \\ &\quad + \varepsilon^2 \left(\frac{1}{96} (21x^2 - 3x^4) \cos(x) + \frac{1}{96} (-21x + 10x^3) \right) \end{aligned}$$

$$\sin(x) \Big) + \varepsilon^3 \left(\left(\frac{35}{192}x^3 - \frac{35}{128}x - \frac{7}{384}x^5 \right) \cos(x) \right. \\ \left. + \left(\frac{35}{128} + \frac{7}{96}x^4 - \frac{35}{128}x^2 - \frac{1}{384}x^6 \right) \sin(x) \right) + O(\varepsilon^4)$$

However, inspection of this solution reveals that the residual of the truncated series is

$$\Delta = \frac{1}{384}x(-105\sin x - 70x^3\cos x - 28x^4\sin x \\ + 105x\cos x + 105x^2\sin x + 6x^5\cos x \\ + x^6\sin x)\varepsilon^4 + O(\varepsilon^5)$$

which contains a term $\varepsilon^4 x^7 \cos(x)$ when expanded. Since $z'' + (1 - \varepsilon x)z - \Delta = 0$, this is a *small* perturbation of the original equation only when $\varepsilon^4 x^7 \cos(x)$ is small compared to $-\varepsilon x \cos(x)$; hence $\varepsilon^3 x^6 \ll 1$ or $x \ll \varepsilon^{1/2}$. However, outside of this interval, the effects of the perturbation can be arbitrarily large, and so the problem would be considered singular in that region.

By using various methods that consist in rescaling the problem to eliminate the secular terms, such as Lindstedt's method, renormalization, or multiple scales, we can find an improved solution which is hardly more complicated,

$$z = e^{\varepsilon x/4} \cos\left(x - \frac{\varepsilon t^2}{4}\right) + O(\varepsilon^2), \quad (46)$$

and it is valid for $0 \leq x \leq O(1/\varepsilon)$. Since $\varepsilon^{-1/2} \ll 1/\varepsilon$, this is an improvement. Some applied mathematicians think of these methods as being appropriate for 'singular' perturbations, and consider this and problems like it to be 'singular'. After all, the residual with the result from the regular perturbation method is not uniformly small as $\varepsilon \rightarrow 0$; it is only small on intervals not as large as $O(\varepsilon^{-1/2})$. Notice also that if $|x| > 1$, then the perturbation is itself not small.

Limitations

Perturbation methods are extremely powerful and apply in an incredibly varied array of situations. Yet, they have some limitations. A basic limitation of perturbations methods is that they

do not apply to discontinuous problems; at least, not unless they have been regularized (transformed into smoother problems). Also, the cost of computing the series may grow quite rapidly with the order of the expansion. We are aware of cases with cost growing exponentially with N if done naively; in many cases, this can be improved to polynomial cost, such as $O(N^4)$, but even so this can be expensive.

Perhaps more interestingly, the results are valid only in a (possibly very small) region about the expansion point ($\varepsilon = 0$); this can give sometimes only a very limited view of the full range of behavior of the model. But the main limitation is the fact that to find p_0 requires solving a nonlinear problem. In general this is always the hardest part. To use perturbation methods, one has to have an initial approximation $p_0 \phi_0(\varepsilon)$; the problem of finding each successive p_k for $k \geq 1$ thereafter is (usually) linear. These two limitations may be addressed separately using a technique known as *homotopy*, or continuation. This can be a numerical technique but is not necessarily so, and it is helpful in a broad collection of situations (see, e.g., Sommesse & Wampler, 2005)

Nicolas Fillion and Robert M. Corless

See also Complex Systems; Kinetic (Molecular) Theory; Physics, Contemporary

Further Readings

- Batterman, R. W. (2002). *The devil in the details: Asymptotic reasoning in explanation, reduction, and emergence*. Oxford University Press.
- Bell, J. L. (2005). *The continuous and the infinitesimal in mathematics and philosophy*. Polimetrica.
- Bell, J. L. (2008). *A primer of infinitesimal analysis* (2nd ed). Cambridge University Press.
- Bender, C., & Orszag, S. (1978). *Advanced mathematical methods for scientists and engineers: Asymptotic methods and perturbation theory* (Vol. 1). Springer Verlag.
- Boyer, C. B. (1949). *The history of the calculus and its conceptual development*. Dover Publications.
- Corless, R. M., & Fillion, N. (2013). *A graduate introduction to numerical methods, from the viewpoint of backward error analysis*. Springer.
- Corless, R. M., Gonnet, G., Hare, D., Jeffrey, D., & Knuth, D. E. (1996). On the Lambert W function. *Advances in Computational Mathematics*, 5(1), 329–359.

- Hoffman, P. (1998). *The man who loved only numbers: The story of Paul Erdős and the search for mathematical truth*. Hyperion.
- Jeffreys, H. (1962). *Asymptotic approximations*. Oxford University Press.
- Robinson, A. (1966). *Non-standard analysis*. Amsterdam, the Netherlands: North-Holland Publishing Company.
- Simmonds, J. G., & Mann, J. E. (1998). *A first look at perturbation theory* (2nd ed.). Dover Publications.
- Sommese, A. J., & Wampler, C. W. (2005). *The numerical solution of systems of polynomials arising in engineering and science* (Vol. 99). World Scientific.

Notes

1. The importance of this distinction has become an important topic of discussion in philosophy of science (see, e.g., Batterman, 2002).
2. This example illustrates convergence. Archimedes in fact demonstrated more, namely the existence of the unique number that we call π , by demonstrating that the perimeters of inscribed polygons ended to the same number. We could have used this example to illustrate perturbation theory itself, by noting that each p_k is a constructible perimeter of a perturbed figure.
3. For sequences of functions, one has to be concerned with the convergence being *uniform*; we will keep discussion of uniform convergence to a minimum but note its importance both in theory and in practice.
4. See Boyer (1949) for a nice history of calculus and Bell (2005) for a history of infinitesimals. It must be noted that many new approaches have revived the use of infinitesimals, such as nonstandard analysis (see, e.g., Robinson, 1966) and smooth infinitesimal analysis (see Bell, 2008).
5. Expansion at infinity is scarcely more difficult. One uses gauge functions that vanish ever more rapidly at infinity, and one typically uses another letter (say t).
6. Because asymptotic power series are unique, an asymptotic power series equal to zero must have each coefficients equal to 0 (see, e.g. Jeffreys, 1962, p. 6). This fact justifies Equations 24 and 25. Note in particular that convergence as $N \rightarrow \infty$ is not needed.
7. We borrow this example from Bender and Orszag (1978, p. 327).

PHENOMENALISM

Two principal schools of thought, empiricism and rationalism, arose out of the philosophical tradition of the Enlightenment period, and of these two

schools, phenomenism is directly derived from, and inspired by, empiricist theories of reality. While there are many nuances and variations in phenomenalist thought, phenomenism can be fundamentally described as a set of theories that maintain that an individual's experience of reality is not comprised of their direct apprehension of real, physical objects but is in fact comprised solely of sensory stimuli, or, *phenomena*, that cannot guarantee any knowledge or certainty with regard to a real and physical world.

This radically empiricist notion that the mind experiences reality only as a collection of sensory perceptions is to be sharply contrasted with the opposing theories of the rationalist Enlightenment philosophers, like Benedict (Baruch) Spinoza and Gottfried Leibniz, who held that reason, or thought, could be used as a tool to legitimately ascertain certain truths and facts about a real, physical world. The first modern theories of phenomenism were advanced by the Irish empiricist George Berkeley, which were then further developed and principally modified by later philosophers such as David Hume, Immanuel Kant, John Stuart Mill, Ernst Mach, and Bertrand Russell.

While all of the aforementioned philosophers deal with phenomenist theory in their respective philosophical works, they do so in different ways, to different extents, and with different results. Some of the theories of the phenomenist philosophers have theological implications or underpinnings, like those of George Berkeley, while others are based upon nontheological and more scientific worldviews, like those of Ernst Mach.

The Theorists and Their Concepts

George Berkeley

In contemporary circles, the Irishman George Berkeley (1685–1753) is regarded as one of the fundamental modern initiators of phenomenist thought. Berkeley's phenomenism is developed by his two main philosophical foundations: idealism and immaterialism. His idealism is based in his belief that physical objects are *mind-dependent* and exist only as an assemblage of ideas, while his immaterialism is based on his belief that there is no such thing as physical substances and that only mental substances exist. Of these mental substances, to him there are two: a finite mental substance and an infinite mental substance constituted by a divine being, God.

These notions form the basis for Berkeley's version of phenomenalism, which he termed *subjective idealism*, and which he employed to negate any certainty regarding the objective existence of objects. He accomplished this by way of declaring that the ideas within a mind are the sole things that can be understood as existing or having reality and that any knowledge or facts garnered about entities outside of the mind are baseless and without merit. Furthermore, Berkeley takes his immaterialism so far, in fact, that he maintains that when a sensory stimulus is no longer perceived, it no longer exists except in the infinite mental substance, or mind, of God. As mentioned earlier, this is the theological implication of Berkeley's phenomenalism, and considering his position as the Bishop of Cloyne, it is an idiosyncratic coalescing of his philosophical and theological concerns.

David Hume

The British-born David Hume (1711–1776) is widely regarded as one of the preeminent philosophers of the Western philosophical tradition, and his work, which dealt with phenomenalism, helped to cement his standing as one of the most distinguished philosophers of the empiricist philosophical tradition. Hume maintained that it is impossible to intellectually move beyond or outside of experience, which itself was merely composed of perceptions. Such intellectual conditions, in Hume's view, would promote only illusions as opposed to certainties. Likewise, Hume enlists his personal understanding of perceptions to negate the philosopher John Locke's conception of material substances, in that Hume claims they are only perceptions. Moreover, Hume critiques George Berkeley's phenomenalism by maintaining that any Berkeleyan examination of mental substances can only lead back to perceptions.

Hume uses his phenomenalism, based upon perceptions, as the keystone to the entirety of his philosophical theories. For instance, Hume's scathing critique on the law of causality is centered on his phenomenalism; he maintains that the rationalistic interpretation that every event must have a cause cannot support consequent judgments, since humans can only perceive perceptions, thus, they cannot piece together the necessary connection between an event and its cause. To Hume, the mind can only be

categorized as an assortment of perceptions, and therefore, there can be no rational knowledge garnered from outside of such perceptions. Hume's phenomenalism allowed him to examine how relations of ideas can be logically true, yet totally extraneous with regard to reality, and it allowed him to explain that individuals may gain knowledge through sensory experience, but that since every individual's experience would be different, one could gain knowledge only of perceptions.

Immanuel Kant

Born in Königsberg, Prussia (now Kaliningrad, Russia), Immanuel Kant (1724–1804) is nearly universally acclaimed as one of the most important philosophers of the modern era. Kant's great intellectual achievement was his ability to synthesize the two dominant philosophical traditions of his time, empiricism and rationalism, into his own personally devised *transcendental idealism*, which sought to utilize and simultaneously transcend both traditions. Kant's phenomenalism differs from Berkeley's stringent idealism in that he believes that *appearances*, or phenomena, are derived from, and point toward, the unknowable *things-in-themselves*, or *noumena*. As detailed in Kant's greatest work, *The Critique of Pure Reason*, Kant critiques Berkeley's idealism further by claiming that when one perceives something, it is a perception of an object as it appears to us, and not merely an idea.

Furthermore, Kant's phenomenalism is quite distinct from Berkeley's and Hume's phenomenalism in that Kant believes the temporal succession of phenomenal appearances is absolutely indicative of a noumenal, though unknowable, reality. Whereas Berkeley and Hume would negate such a notion by saying that an individual can only perceive their subjective ideas or perceptions, Kant believes that we can in fact perceive objects but only in their phenomenal appearances. According to Kant, there must be some noumenal reality that orders the phenomenal reality or else no phenomenal reality could appear. In this way, he is quite distinct from his other philosophical contemporaries who dealt with phenomenalism.

John Stuart Mill

John Stuart Mill (1806–1873) was a prodigious British philosopher who is critically acclaimed as

one of the most significant thinkers of the 19th century. While Mill studied and wrote about and worked on a wide array of philosophical and social issues, his philosophical foundations were strongly empiricist, and thus, his theories on phenomenalism arise out of his intellectual empiricism. Consequently, Mill believed that when an individual interacts with a physical object, they are subjected to an array of sensory stimuli. Furthermore, such interactions with physical objects can produce different sensations under different circumstances. Mill claimed that after repeated interactions with particular objects, we as humans come to expect certain sequences of sensations under certain circumstances. Thus, Mill's phenomenalism boils down to his notion that physical objects are composed of their consistent possibilities of sensation.

In other words, to Mill, physical objects are assortments of sensations that could, will, and do occur upon interaction. From the complex network of these sensations and their possibilities are derived the concepts of spatial dimensions, viewpoints, and other similar notions. So, in a sense, Mill's phenomenalism is as contingent upon experienced phenomena as it is upon the unexperienced possibilities of phenomena. Mill believed that the actualities and possibilities of phenomena are lawfully and regularly ordered, and thus, we can have provisional certainties regarding their occurrences; from these actualities and possibilities, we take collections of ideas and then construct inferences concerning our interactions with physical objects to the point that certain inferences become nearly unconsciously automatic.

Ernst Mach

Ernst Mach (1838–1916) was an Austrian theorist who made notable contributions to the fields of philosophy, physics, and psychology. Mach's phenomenalism is different from that of the previous phenomenalist theorists mentioned, in that Mach derives his phenomenalism more from his scientific and psychological understanding than from any philosophical position. Still, the result is similar, for Mach was supremely interested in how the physical body digests, or processes, sensory stimuli. Through an intensive study of the human retina, Mach deduced that the mind does not apprehend stimuli directly but

rather through an embodied network of neural relations. As a consequence, Mach maintains that one does not experience reality firsthand but only its aftereffects that result from our nervous system's processing of it.

Thus, Mach's phenomenalism can be described by his belief that we do not encounter reality directly but rather through the organic construction of it as a result of the neurologic processing of sensory data. In this way, it can be seen how Mach's scientific and psychological background informs his phenomenalist conception of the mind. An important result of Mach's phenomenalism is that it seems to negate philosophical representationalism, in that it seems to make a straight bridge between appearance and reality unsustainable.

Bertrand Russell

The British philosopher Bertrand Russell (1872–1970) is regarded as one of the principal initiators of the analytic philosophical tradition and one of the most important thinkers of the 20th century. Russell's phenomenalist theories are similar to John Stuart Mill's phenomenalism, in the sense that Russell believed that our perceptions of sensory stimuli are not produced by physical objects but instead actually comprise them. In this way, to Russell, physical objects *are* sensory stimuli and not simply something that causes them. The difference between Russell's and George Berkeley's conceptions of phenomenalism, respectively, is that Russell believed that if a consciousness is removed from a particular set of sensory stimuli, there would still be a perspective from which to perceive such stimuli, whereas in a similar situation, Berkeley would maintain that the stimuli all but ceases to exist except in the mind of a divine being.

To describe this phenomenalist conception of reality, Russell utilizes the terms *private space* and *perspective space*. Private space can be designated as the perspective in which a single consciousness comes to perceive sensory stimuli, while perspective space can be defined as the totality of perspectives in which a multitude of private spaces interact and intersect. For Russell, different perspectives provide different aspects of the sensory stimuli that comprise physical objects.

William M. Peaster

See also Concept; Epistemology; Philosophy of Mind

Further Readings

- Anderson, C. A., & Savage, C. W. (Eds.). (1989). *Rereading Russell: Essays on Bertrand Russell's metaphysics and epistemology*. University of Minnesota Press.
- Berkeley, G. (1993). *Philosophical works, including the works on vision*. J. M. Dent & Sons.
- Bird, G. (1962). *Kant's theory of knowledge: An outline of one central argument in the Critique of Pure Reason*. Routledge & Kegan Paul.
- Breck, A. D., Yourgrau, W., & Kaplan, D. (Eds.). (1970). *What is Russell's theory of descriptions?* Anchor Books.
- Ferrier, J. F. (2007). Dissolving Hume's paradox: On knowledge of mind and self. *Journal of Scottish Philosophy*, 5(1), 7–13. <https://doi.org/10.3366/JSP.2007.5.1.7>
- Guyer, P. (2006). *Kant's Copernican revolution: The refutation of idealism*. Routledge.
- Harris, J. A. (2005). *Of liberty and necessity: The free will debate in eighteenth century British philosophy*. Clarendon Press.
- Kolak, D., & Thomson, G. (2006). *The Longman standard history of modern philosophy*. Pearson Education.
- Laine, M. (Ed.). (2001). *A cultivated mind: Essays on John Stuart Mill presented to John Robson*. University of Toronto Press.
- Letwin, S. (1965). In Michael Ayers and Daniel Garber, *The pursuit of certainty: David Hume, Jeremy Bentham, John Stuart Mill, Beatrice Webb* (pp. 513–552). Cambridge University Press.
- Longuenesse, B. (1998). *Kant and the capacity to judge: Sensibility and discursivity in the transcendental analytic of the Critique of Pure Reason*. Princeton University Press.
- Miller, L. (2009). *Questions that matter: An invitation to philosophy*. McGraw-Hill.
- Nadler, S. (1998). Doctrines of explanation in late scholasticism and in the mechanical philosophy. In J. Barwise, H. J. Keisler, P. Suppes, A. S. Troelstra, *The Cambridge history of seventeenth-century philosophy* (pp. 641–660). Cambridge University Press.
- Steen, I. K. (2004). *Russell on matter and our knowledge of the external world*. Syracuse University Press.
- Wolters, G. (1989, April). Phenomenalism, relativity and atoms: Rehabilitating Ernst Mach's Philosophy of Science. In *Logic, methodology and philosophy of*

science VIII: Proceedings of the Eighth International Congress of Logic, Methodology and Philosophy of Science, Moscow, 1987. Elsevier Scientific. (Original work published 1987)

PHILOSOPHY OF MIND

Philosophy of mind is a branch of philosophy concerned with the nature of mental events and processes. Inquiries into philosophy of mind can be traced back to ancient Greece, whereas interest in the nature of the mind can be found in many of the world's religions. Contemporary philosophy of mind is largely an interdisciplinary effort. Many contemporary philosophers of mind draw from work in computer science, neuroscience, psychology, biology, sociology, physics, and chemistry. However, a growing number of contemporary philosophers of mind have begun to borrow insights into the mental from other branches of philosophy such as phenomenology, existentialism, pragmatism, and postmodernism.

Contemporary work in the philosophy of mind varies in scope. As will be seen in what follows, many contemporary philosophers of mind work on local issues of mind such as the nature of mental content, the nature of the self, free will, emotion theory, and perception. Other contemporary philosophers of mind work on the nature of consciousness, the place of consciousness in nature, and the relationship between consciousness and the body.

Approaches to the Metaphysics of Mental Phenomena

At the most global level and in its most common manifestation, philosophy of mind attempts to provide a framework for understanding the place of consciousness in nature, the relationship between mental processes and other natural processes, and the nature of mental phenomena in general. Often, research on these issues is conducted within a particular theoretical paradigm that contains several key tenets, or positions, on the nature of mental phenomena. For most of these paradigms, the salient tenets of that paradigm

are derived from a number of key methodological and other metaphysical commitments. The following sections summarize many of the prevailing perspectives on the metaphysics of mental phenomena.

Brain State–Mental State Identity Theory

The *brain state–mental state identity* (BSMSI) theory maintains that, although perhaps conceptually distinct, mental states, or processes, are identical to states, or processes, in the brain. As an analogy for their position, early advocates of the BSMSI, or identity, theory evoked the relationship between lightning and an electrical discharge. Just as lightning is nothing more than an electrical discharge, the advocate of the BSMSI theory held that conscious phenomena are nothing more than various brain processes. Early advocates of the view hoped to clarify their position by drafting a distinction between *meaning* and *reference*. According to these thinkers, two terms may differ in terms of meaning, but not differ in terms of reference. On this view, the difference between mental states and brain states could be mapped onto a difference in the meaning of the terms, even though the terms shared a referent. So, in the case of lightning and electrical discharge, lightning and electrical discharge have different meanings even though they refer to the same phenomena.

Most advocates of the identity theory endorse methodological reductionism, understood as the collapse of one set of theoretical terms into another. To accomplish this reduction, proponents of this type engage in a three-part process. In the first part, the proponent seeks to clarify, or individuate, terms from folk psychological language (consciousness as it is described by the *folk*). In the second part, the proponent seeks to clarify, or individuate, terms from various brain sciences. In the third part, the proponent seeks to fix references between the two terms. Successful reference fixing should mean that the terms taken from the folk psychological language and the relations between them may be describable solely in the terms and relations provided by the brain sciences.

Mental State Eliminativism

The eliminativist's view of mental processes shares much in common with the proponent of identity theory insofar as the eliminativist views the mental processes as nothing other than various physical processes of the brain. Where the eliminativist makes breaks from the identity theorist is in the disavowal of folk psychology entirely. While the identity theorist retains folk psychology as a critical component of the reference-fixing process, the eliminativist maintains that folk psychology is so deeply erroneous that the inclusion of folk psychological terminology would not only hinder the development of a successful theory of mind but also point the theorist in the entirely wrong direction. That is, the eliminativist not only maintains that folk psychology is deeply flawed, but that folk psychology is mistaken about the kinds of mental phenomena that there are. So, for example, folk psychology generally progresses on the assumption that things like beliefs, emotions, desires, and so on are posited as true mental phenomena. The eliminativist may hold that none of these phenomena actually exists, and thus that any attempt to make a biological account of brain processes cohere with these phenomena is deeply misguided. Of course, it is important to note that a number of theorists, including those who would not describe themselves as eliminativists, challenge certain assumptions of folk psychology. The difference between these individuals and full-fledged eliminativism is a matter of degree.

Because the eliminativist disavows folk psychology entirely, the eliminativist cannot judge, as the identity theorist can, the success of a project on the basis of its reference-fixing success. Here, the eliminativist is aided by the methodological commitments of a particular scientific program. That is, to judge the success of an account, the eliminativist need only appeal to the demands of its theoretical paradigm. So, for example, an eliminativist constructing a neurobiological account of some particular phenomena need only judge the success of such an account in regard to the theoretical desiderata of neurobiology. If the account meets the desiderata, then the eliminativist can consider such an account successful. If the

account does not meet the desiderata, then the eliminativist must deem such an account unsuccessful. Again, the critical point of emphasis is that the eliminativist need not judge the success of an account on its ability to accommodate folk psychological phenomena such as beliefs or desires.

Functionalism

Functionalism is the view that a mental state, or process, cannot be defined in terms of that state's constitution but must be defined in terms of the role it plays within the overall system in which it operates. Typically, functionalists take a mental state, or process, to be determined in virtue of its standing in specific causal relations between itself, sensory information, other mental states, and behaviors. Functionalism, so construed, supports the intuition that constitutionally diverse systems may instantiate, or realize, identical mental states. So, for a simplified example, a functionalist may conceive pain as having the function of producing agitated states, be they agitated mental states or physical behaviors, and thus conceive both humans and canines as capable of being in the mental state pain without having to posit underlying identical substrates. That is, a human and a dog may both be in pain, even though a human's pain may be caused by the firing of C-fibers, and dog pain may be caused by some other source.

Although functionalism was formulated and defended by thinkers as early as the Scholastics, contemporary functionalism is very much indebted to the work of the English mathematician and computer scientist Alan Turing (1912–1954). On Turing's view, intelligence could be measured against other assumed intelligences, such as the intelligence of a human, by determining whether a particular system was capable of producing intelligible responses to particular queries formulated by the assumed intelligence. Today, this sort of measure is known as a *Turing test*, but its crucial contribution to philosophy of mind was the identification of thoughts with various functional roles, which, in the case of the Turing test, would be the system's ability to produce other internal states and appropriate responses.

The contemporary equivalent of Turing's view is called *machine state functionalism*. On this view, any minded creature's operations can be fully specified by the following instruction: If a machine is in state S_i and receives input I_j , then the machine will enter state S_k and produce output O_l . However, because the machine state view equates a mental state with the *total* state of the system, and many assumed intelligent systems, such as humans, appear capable of supporting complex but distinct internal states, many contemporary functionalists opt for another variety of functionalism.

One particularly popular contemporary view is sometimes referred to as *psycho-functionalism*. For the proponent of *psycho-functionalism*, the main point of departure between this view and the machine state functionalist's view is the psycho-functionalism's disavowal of Turing's instruction. Instead, the psycho-functionalism adopts the methodology of cognitive psychology and instead characterizes mental states in virtue of their role in a particular psychological theory. So, on the view of psycho-functionalism, a particular organism is in state M , if M is specified by some theory T and M plays functional role F as specified by T . However, many contemporary functionalists now defend carefully nuanced views that may not be easily described as either machine state functionalism or psycho-functionalism.

Embodied Cognition

Embodied cognition is the view that the mental states, or processes, of an agent are both causally and constitutionally dependent on that agent's bodily characteristics. Proponents of *embodied cognition* may be placed on a continuum of decreasing emphasis on the brain in cognition such that proponents of another approach to the study of mental phenomena may also be considered to be, at least in part, dedicated to the embodied cognition research program. Nevertheless, the majority of researchers self-described as proponents of embodied cognition typically also perceive themselves as combating the notion that cognitive processing may be described in abstraction from bodily mechanisms. These thinkers

stress the use of the environment by an agent to abet cognitive processing and the informational role of bodily feelings in cognitive processing while maintaining that phenomenal consciousness of an agent, what it is like to be that agent, does not extend beyond the skin, or other theoretical border, of the agent.

Although embodied cognition has generated significant interest in the past two decades, traces of the view can be found in St. Thomas Aquinas. On Aquinas's view, the soul, or mind, of an agent is considered as the form of that being. Contemporary advocates of embodied cognition share the monist understanding of the relationship between the mental and the physical with Aquinas and generally market their view as a radical departure from the homuncular understanding of the mind proponents of this view charge dualists, functionalists, and identity theorists with employing.

Typically, proponents of embodied cognition stress the importance of sensory-motor skills in accomplishing various cognitive feats such that an agent is thought to structure their world. Thus, proponents of embodied cognition challenge the division other approaches to the study of mental phenomena construct between the features of the external world the representation of those features internal to the agent in virtue of the division's inability to accommodate the extensive feedback loop that exists between the agent-world coupling. In the view of embodied cognition, knowledge emerges in the dynamic interaction between an agent and the environment such that an agent creates its world through its embodied actions. One crucial implication of embodied cognition, then, is that whether or not an agent possesses certain cognitive capacities depends on the material constitution of that agent. On this view, an agent having hands, legs, and eyes, as well as the skill to coordinate these body parts, may very well determine the kinds of intellectual feats that the agent is capable of.

Extended Mind Theory

Extended mind theory shares much in common with several key tenets of the embodied

cognition view. Where the extended mind theory deviates from embodied cognition is in its commitment to the position that the phenomenal consciousness of an agent, what it is like to be that agent, may extend beyond the boundaries of that agent's body. However, this is not to suggest that proponents of the extended mind theory do not recognize the importance of an agent's body in cognition. Rather, many proponents of extended mind theory often argue that a proper understanding of the dynamic interaction between the agent and the environment points not just to the recognition of the importance of the agent's body in cognition but even further to the recognition that cognition cannot be described as either internal or external to the agent. For example, embodied cognition theorists often defend their position through an appeal to the apparent coupling between visual perception systems and other sensorimotor capacities in particular environments. Extended mind theorists argue that coupling between the visual systems and other sensorimotor capacities, which the embodied cognition theorist takes as indicative of their both sharing in the constitution of the mind, is not much different from the coupling between the agent's body and the environment. If coupling indicates shared constitution, then the extended mind theorist argues that the environment must play a constitutional role in an individual's mind as well.

Panexperientialism

Panexperientialism is the view that all real relations in the world may, at some level, be described as mental properties. Panexperientialism is to be distinguished from the view of naive panpsychism, which holds that every real entity has a mind. Rather the panexperientialist maintains that the ultimate constituents of reality are most appropriately described in terms of phenomenal, that is, *felt* sensations. The panexperientialist conceives more complex-minded creatures, such as animals, as particularly arranged collections, or loci, of these phenomenal properties. Importantly, panexperientialists differ as to how these fundamental properties may be arranged and how their

arrangement may account for the apparent variation in differently minded creatures.

Although panexperientialism is a minority view in both the philosophical and scientific communities, the view has precedent in philosophical history. A version of panpsychism was defended in the early modern period by Gottfried W. Leibniz (1646–1716), and more recent views have been defended by so-called *process philosophers*, the most notable of which are Charles Hartsorne (1897–2000) and Alfred North Whitehead (1861–1947). While the latter two hold distinct views from Leibniz, all three employ a similar and commonly evoked argument in support of their view. On this account, the macroscopic world is a sort of record of a microscopic reality such that any explanation given in the microscopic reality must be illustrated in the macroscopic world. At some point, these thinkers claim, the explanation must bottom out and the question becomes one concerning the nature of what we might call ultimate individuals. These thinkers' answer is that the illustration we must find in the macroscopic world need be neither visual nor tactile. Rather the answer itself is given in experience. At any moment in time, our experience reveals itself as individual; anything which is an object of awareness is always presented to an individual. Thus, the nature of the ultimates, panpsychists argue, is experiential.

Other panexperientialists argue along more theoretical, but similar, lines. In this view, mental properties are identified as intrinsic, or nondispositional, properties, while the properties of contemporary science are identified as relational, or dispositional, properties. Given the conviction that relational properties cannot exhaust the set of real properties, the panexperientialist argues that the nondispositional properties which enter into the relations identified by contemporary science must be the phenomenal properties the panexperientialist claims that experience tells us are intrinsic. Thus, for these theorists, a proper account of more complex cognition, such as human cognition, must cut across many (if not all) of the domains of contemporary science. However, it is worth noting that panexperientialists may support their arguments in a number of other ways and incorporate elements of many of the other approaches to the study of mental phenomena.

Neutral Monism

Neutral monism is the view that there are intrinsic, nondispositional, fundamental properties in reality and that these properties may enter into relations, which may be described as either mental or physical. An early version of this view was defended by the Dutch philosopher Baruch Spinoza (1632–1677). In Spinoza's view, both an agent's mind and body are manifestations of an underlying neutral substance. In the early 20th century, views of this sort were defended by both William James (1842–1910) and Bertrand Russell (1872–1970). Although the view shares much in common with panexperientialism, neutral monism may be distinguished from panexperientialism insofar as neutral monism sees the fundamental constituents of reality as neutral, while the panexperientialist sees the fundamental constituents of reality as primitive mental properties.

Dualism

Dualism is the view that mental phenomena are a distinct type of thing from other physical phenomena. Typically, the view is formulated on several key distinctions, such as the distinction between observable and unobservable properties, intentionality and nonintentionality (i.e., whether an object is about something); extension and nonextension, where mental phenomena are considered unobservable, intentional; and nonextended and physical phenomena are considered observable, nonintentional, and extended. (It should be stressed, however, that some dualists deny mental phenomena must necessarily have each of these properties.)

Dualism has a long history in the philosophical canon. An early example of the view was defended by the ancient Greek thinker Plato in the *Phaedo*. There, Plato argued that the soul must be immaterial from his doctrine of Forms, which he contended not only govern all of reality but render it intelligible and is itself immaterial. Assuming that to perceive a Form, the intellect must share something in common with the Form, Plato argued the soul must be itself immaterial. Some scholars attribute a similar argument to Aristotle. According to these scholars, in the *De Anima*, Aristotle argued that the soul must be immaterial because it has no restricted range of objects which it may consider, and every faculty

which corresponds to a material organ, such as sight, sound, and touch, have a restricted range of objects they can receive.

However, the most common form of dualism, *substance dualism*, is first attributed to the early modern French philosopher René Descartes (1596–1650). In this view, mental phenomena are properties of a substance, distinct from the physical properties of a human's body. Moreover, this view generally maintains that a person is to be identified with the mental substance rather than the physical substance or the joint mental–physical entity.

In the modern period following Descartes, and even to this day, dualism has generally been dismissed on the basis of its inability to give an account of how two distinct substances can interact. Some contemporary proponents of the substance view of dualism counter this objection with the reply that one cannot rule out a priori the interaction of two distinct substances. They accept the interaction objection and instead opt for a more naturalistic version of dualism, which has been proffered under the name *double aspect theory* or *property dualism*.

In the double aspect view, mental and physical properties are different properties of one underlying fundamental type of entity. Although double aspect theory shares much in common with neutral monism, the view is to be distinguished from the latter in that neutral monism allows for fundamental constituents to express themselves as mental, physical, both, or neither while double aspect theory maintains that the mental and physical are inseparable, irreducible, but necessary properties of reality.

Despite the double aspect theorists' naturalistic take on dualism and their abandonment of the notion of distinct types of substances, double aspect theory has faced similar interactionist challenges as substance dualism. Like the contemporary substance dualists, some double aspect theorists object to the interactionist challenge. Others, however, accept the challenge and deny that such interaction takes place. One such group, the *epiphenomenalists*, maintain that the mental states of an individual have no bearing on the physical states of that person whatsoever and are merely *along for the ride*, in a manner analogous to the way the shadow cast by a speeding car in the sun is *along for the ride* with the speeding car.

Another group, *parallelists*, maintain that causation occurs within property types and that the parallel causation within each type merely runs parallel with the other type.

Central Issues in the Metaphysics of Mental Phenomena

Because interest in the nature of mental phenomena has permeated philosophy for centuries, many philosophers of mind have aimed not only to provide arguments in favor of their particular metaphysical stance but also to provide arguments against other prevailing metaphysical stances. Importantly, these negative arguments may be directed at a stance distinct from one's own, or at a stance similar to one's own. For example, a functionalist may not object only to dualism but also to other functionalist programs. The following is a brief presentation of some of the central issues debated in contemporary discussion on the metaphysics of mental phenomena.

Qualia

The term *qualia* has appeared in discussions on philosophy of mind for almost a century. It was first used by C. I. Lewis (1883–1964) in 1929 in a discussion of the sense-datum theory of perception. However, its contemporary use has been extended to refer to the properties of phenomenal experience, what it is like to have that experience, altogether. A singular property of a synchronic moment is called a *quale* and a number of these properties are called *qualia*. So, for example, the sight of a red book would involve one quale, what it is like to see the red book, and the sight of a blue book would involve another quale, what it is like to see the blue book, while both sensations are dubbed qualia.

At this level of description, the existence of qualia is uncontroversial. However, a number of contemporary philosophers have contended that qualia have the additional features of being non-representational, intrinsic properties of an experience. According to these theorists, qualia are accessible to introspection, mental properties corresponding to some salient property of an object, the sole determinant of the character of an

experience, and not covarying the representational content of the object to which the quale corresponds. Whether or not qualia truly possess these features, and to which mental states qualia correspond, has generated significant debate.

Within the discussion of qualia, there has been the debate over whether or not qualia are reducible to other physical properties. This issue has been of particular importance to the philosophy of mind because if qualia are not reducible to other physical properties, then the reductive approaches to the study of mental phenomena, such as the mental state–brain state identity theory, mental state eliminativism, and some versions of functionalism, must be flawed and incorrect. Central to this debate are a number of key characters and thought experiments including color-blind neuroscientists and philosophical zombies.

For example, some philosophers of mind have argued that a brilliant but color-blind neuroscientist with complete, exhaustive knowledge of the neurobiological mechanics of color perception would not know what it is like to see color, a phenomenon. Such philosophers further suggest that knowledge of physical truths therefore does not exhaust the knowledge of mental truths and thus that mental properties cannot be reduced to physical properties. Similarly, some philosophers of mind have argued that creatures physically identical to humans, but without phenomenal experience (philosophical zombies), are conceivable and that their conceivability demonstrates the irreducibility of mental phenomena to physical properties.

Content Internalism and Externalism

With the increasing popularity of the situated views of mental phenomena (embodied cognition and the extended mind theory and various versions of functionalism), contemporary philosophers of mind have become increasingly interested in the extent to which cognition represents an agent's world and how these representations gain their meaning or content.

According to *content externalists*, the meaning of an individual's concept is determined by the relationship between the individual and the

individual's environment. A classic thought experiment in support of this position considers two individuals whose mothers are identical twins and who are raised in isolation from one another, but in similar environments. When each individual expresses the statement "I want my Mom," the individual's wish will be satisfied only if the individual is brought the individual's mother and not an identical sister, even if the individual child does not know that the wish is not satisfied. So, the externalist argues, if individuals know what they mean and say what they mean, the difference between each individual's thought "I want my Mom" may not amount to a difference in the internal constitution of either individual.

Content internalists, on the other hand, maintain that the meaning, or content, of an individual's concept is determined by the various details of the individual's internal constitution. A content internalist maintains that the content of an individual's thoughts can be discerned by carefully observing the subject's mind, or if the content internalist endorses a reductionist view of mental phenomena, the substrate of the subject's mind. Other philosophers of mind argue that meaning is determined by both internal and external factors. For example, some philosophers of mind recommend two-dimensional semantics, which takes each concept to have both an external (sometimes called a *wide content*) component and an internal (sometimes called a *narrow content*) component.

Mark D. Sherry and
Walter Scott Stepanenko

See also Computer Science; Life Sciences, Contemporary; Metaphysics; Neuroscience; Phenomenalism

Further Readings

- Block, N., Flanagan, O., & Guzeldere, G. (Eds.). (1997). *The nature of consciousness*. MIT Press.
- Chalmers, D. (1996). *The conscious mind*. Oxford University Press.
- Churchland, P. M. (1989). *A neurocomputational perspective*. MIT Press.
- Clark, A. (2008). *Supersizing the mind: Embodiment, action, and cognitive extension*. Oxford University Press.

- Kim, J. (1998). *Mind in a physical world*. Bradford.
- Noë, A. (2004). *Action in perception*. MIT Press.
- Skrbina, D. (Ed.). (2009). *Mind that abides: Panpsychism in the new millennium*. John Benjamins.
- Smith, B. (2013). Internalism and externalism in the philosophy of mind and language. In P. Saka (Ed.), *The internet encyclopedia of philosophy*.
- Taliaferro, C. (1996). *Consciousness and the mind of God*. Cambridge University Press. ISSN 2161-0002, <https://iep.utm.edu>

PHILOSOPHY OF SCIENCE

All men are mortal.
 Socrates is a man.
 Therefore, Socrates is mortal.

Aristotle's syllogism arguably marks the beginning of formalized scientific reasoning. Linking a general premise ("all men are mortal") to a particular observation ("Socrates is a man") allows for the logical deduction of the conclusion: Socrates is mortal. Reflection on the structure and the methods of scientific reasoning touch on the central questions of the philosophy of science. Although its roots can be traced back to ancient Greece, as a subdiscipline of philosophy, it gained particular prominence only in the late 19th and during the 20th century, when scientific (and technological) progress turned science into the main source of knowledge and a driving force of social change. In the process of scientific development, in particular the formation of modern natural sciences, philosophy and science became more differentiated: While the latter focused on the observation and explanation of natural phenomenon and aimed at the development of theories and discovery of natural laws, the former began to reason on the foundations of scientific explanation. A second philosophical strand emerging as a reaction to the increase in importance of science was the ethics of science. It concentrates on the consequences of the expansion of science and the related technological progress, questioning ethical standards of research and the ramifications and potential dangers of scientific discoveries. In contrast, philosophy of science

focuses on the nature of scientific reasoning: What are scientific theories and methods? By analyzing the methods of inquiry as well as the explicit and implicit assumptions of scientific practice, philosophy of science deals with the fundamental epistemological question of what separates knowledge from belief.

Coming from the Latin word *scientia*, which means knowledge, philosophy of science seeks to answer the question of what scientific knowledge is. In a broad sense, knowledge can be defined as *justified, true belief*. The definition is especially suitable, for it highlights two important aspects of the question about the nature of science: (1) What qualifies a belief to be regarded as true? and (2) How can knowledge claims be justified? Another related, fundamental task is the identification and justification of reliable criteria to demarcate science from nonscience or pseudoscience. Demarcating science from nonscience thus puts philosophers of science in the midst of struggles over the respective authority of scientific and non-scientific knowledge and internal disputes about the relation of the natural and social sciences.

Principles of Scientific Reasoning

Observation

Scientific inquiry is based on observation and the collection of data for formulating and testing hypotheses and finding explanations. Observation can either be direct (x is accessible for measurement) or indirect (existence of x has to be derived from a theoretical framework). Furthermore, phenomena that are *in principle* unobservable (e.g., metaphysical entities) have to be separated from *not yet* observable facts. This basic conviction provides a first, tentative criterion for demarcating scientific from nonscientific knowledge: Science focuses on observables. Unobservable, especially metaphysical, questions, although important for science, are not part of scientific inquiry.

Theories and Methods

Scientific reasoning relies on theories and methods. However, the use of theories and methodological approaches is not sufficient to label a

particular insight scientific. Every observation requires a framework of assumptions to be explainable. The observation that the sun rises every morning in the east and sets every evening in the west, for instance, can give rise to a variety of concepts. One could link the apparent daily “birth” and “death” of the sun to the presence of gods, eventually developing elaborate beliefs, which require specific actions (e.g., prayers, sacrifices) to stabilize this cosmic rhythm. Moreover, theories can be false but still be scientific, while a belief can be true without being scientific. It is therefore necessary to differentiate between the structure of theories and methods on the one hand and the truthfulness of a claim on the other.

What is scientific theory? The classical Greek term *theoria* (θεωρία) literally means “viewing” or “looking at.” A scientific theory is an axiomatic system, which offers a specific point of view and explanation of selected phenomena of nature or the social world. Theories guide scientific inquiry and are important to determine which particular aspects of reality should be observed and analyzed. According to this understanding, scientific theories are not only explanatory models developed to explain observations already made. They represent conceptual frameworks, which help to select and interpret the phenomena that should be observed. Theories are mental models or perspectives, which highlight some characteristics while others are rendered irrelevant for explanation. In general, a theory can be described as a connected set of ideas and propositions. The ideas constitute the concepts (or conceptual ideas) of a theory. These provide an explanatory framework for a defined set of observations and allow the formulations of propositions and hypotheses, which are open for empirical testing. Empirical findings can either verify or falsify a theory. In order to gather such supporting or contradicting empirical evidence, *scientific methods* must be employed. This means that the results and interpretations of an observation must be (1) connected to the hypothesis in a distinct way, (2) comprehensible, and (3) repeatable.

The significance of scientific theories and methods depends on the overall aim of scientific inquiry. A rather simple approach, close to a commonsense understanding of theories, regards them as means to deduct “laws” to explain observed

correlations between particular phenomena. From an instrumentalist’s perspective, theories are more than a means of explanation; they are instruments for making predictions about a particular set of data, that is, a specific fraction of the natural or social world. Theories serve as information processing units. Observable data (input) are processed to generate observable predictions (output). The quality of a given theory in this view depends on its ability to generate predictions regarding what will be observed. The correctness of the predictions will ultimately prove the quality of the theory. Instrumentalism treats scientific theories as means to producing *justified, true belief*. Statements and propositions of a theory are truth-evaluable only if the predictions made on observable data turn out to be correct.

A major criticism of instrumentalism in science is the neglect of impalpable (unobservable) data. Impalpable data must not be juxtaposed with speculative data: Both the force of gravity and the concept of causality are impalpable, yet important for scientific explanation. Moreover, the distinction between observable and unobservable entities is not necessarily sharp or distinct. Improvement in observation techniques and advanced instruments (e.g., microscopes) can shift the boundaries between the two categories. In contrast to instrumentalist approaches, which hide those unobservable entities inside the black box of theories, scientific realism, while emphasizing that there are significant differences in the empirical claims about observable and unobservable data respectively, argue that both have to be considered.

Scientific realism is an umbrella term for positive epistemic approaches toward scientific theories. The overall assumption of scientific realism seems to be rather simple: It is based on the belief in the existence of a real outside world, which exists in a distinct way independently of scientific observation. In contrast to idealist or neo-Kantian approaches, which share the common belief that the experienced reality depends, at least partially, on observation and the mind of the observer (most prominently the Kantian belief, outlined in his famous *Critique of Pure Reason* that *causality* doesn’t exist in the outside world but is a category of human understanding), scientific realists state that theoretical claims themselves constitute knowledge about the world. In this sense, scientific

realism differs sharply from instrumental approaches, for its baseline assumption states that (the best) scientific theories can give true (or approximately) descriptions of the independent, outside world.

The Demarcation Problem of Telling Science From Nonscience and Pseudoscience

Defining scientific knowledge implies to label some knowledge as unscientific. However, this task is notoriously complicated since nonscientific reasoning and even bogus science share some common characteristics with science: Beyond intuitive distinctions between typical scientific fields (e.g., physics, biology, chemistry, or astronomy) and nonscientific fields (e.g., astrology, alchemy, or phrenology) defining satisfying criteria to justify this categorization is controversial. The importance of this problem is underlined by epistemological as well as practical reasons. In contrast to a mere agreement that some specific sets of lore (say, astrology) doesn't qualify for scientificity, philosophy of science raises the claim to explain what makes reasoning scientific. In the history of the philosophy of science, various criteria for the distinction of scientific and nonscientific knowledge have been proposed. A demarcation criterion close to a commonsense understanding of science was introduced by logical positivists such as Ludwig Wittgenstein (1889–1951) and Rudolf Carnap (1891–1970). In their view, scientific statements are meaningful as long and only as they are verifiable. Verification requires the empirical testability of a particular claim.

Karl Popper (1902–1994) suggested an opposing perspective: For him any scientific theory must be falsifiable. His principle of falsification rests on the conviction that the cardinal characteristic of scientific knowledge (which separates it from any other form of knowledge) is the fundamental possibility to prove it wrong. A radical departure from the principle of verification, his approach changed the way that scientific claims have to be formulated. In contrast to belief systems (e.g., religious convictions) as well as most pseudoscientific theories, scientific hypotheses must include clear criteria for their falsification. For instance, asserting a specific relation between an astrological star

sign and a person's character traits is not a scientific hypothesis.

The principles of verification and falsification represent the most common, yet not definitive, criteria for demarcating science from pseudoscience and nonscience. Popper's concept of falsification is widely used to test the scientific character of a given theory. To demand in-principle falsifiability is intuitively a plausible criterion: only theories whose assertions can be tested and falsified can lead to true scientific discoveries. Concepts that lead to general, untestable statements or theories—which can be made to fit any empirical observation—violate Popper's principle of falsifiability. Either the purported link between one's sign of the zodiac and one's personality must be empirically testable or astrology must be labeled as pseudoscience. Likewise, theories that tend to explain away conflicting data by inventing ad hoc explanations to maintain (or simulate) consistency with conflicting data also would not count as scientific. The criterion of falsification can be criticized, however, for being too simplistic. Especially complex theories that include sets of assumptions, hypotheses, and heuristics cannot easily be falsified. In these cases, falsification describes a process rather than a decisive moment when a theory is proven wrong. It is more likely that theories, when continually confronted with conflicting evidence, are running out of explanations. This may result in the *ex post facto* characterization of a theory as pseudoscientific but certainly doesn't necessarily offer the precise demarcation criteria as envisioned by Popper.

Scientific Reasoning: Induction and Deduction

The demarcation criteria of verification and falsification (roughly) correspond with two fundamental approaches of scientific reasoning: induction and deduction. *Inductive reasoning* seeks to derive general principles from individual observation, while the process of *deductive reasoning* starts with some general statements (premises) and aims at assigning logically certain conclusions for individual cases. Induction and deduction employ opposite logical modes of explanation. They have to be further qualified with regard to the principles of verification and

falsification and with respect to universal or probabilistic explanation respectively.

Induction and Hume's Problem

Inductive reasoning starts with specific observation so as to discover general rules or laws, which can be applied to a class of cases of the same kind. Observing a swan and noticing that it is white is one such observation. Additional observations (other swans of the same subspecies) may confirm that mature animals have this distinct color. However, a single observation of a black swan would refute the inductively derived conclusion "all swans are white."

The principles of verification and inductive reasoning are closely related. Linking the meaning of scientific knowledge to its verifiability poses a major problem, for then scientific laws could never be verified: In general, laws in science have the following logical expression:

For all X: if A then B

Observing B confirms A. The connection between A and B is only proven inductively, however, not conclusively. In a logical sense, the confirmation of the interrelation between A and B is necessarily provisional. Scientific laws are not strictly verifiable, for observing B at the time t allows only the confirmation of A_t . Consider the famous anecdote of Newton observing an apple (B) falling from a tree in a straight line. Let's assume that he concludes that the trajectory of the apple is determined by gravity (A) and let us further assume that he presumes some regularity (X). How could he verify his claim that the falling apple "proves" a law of gravity? From the perspective of logical positivism, he could observe more apples drop ($B_{t_1} \dots B_{t_n}$) to confirm the influence of $A_{t_1} \dots A_{t_n}$. In a strict sense, verification of the law of gravity would require the observation of all apples (and for that matter all other bodies) to fall. Arriving at a principle by complete observation is rarely possible. Nevertheless, scientific as well as everyday knowledge often rests on convictions gained inductively. In fact, most of reasoning depends on the assumption that unobserved objects are similar in all relevant aspects to objects

of the same class which have been examined. That means that we assume that our observations made "so far" can be generalized for similar unobserved objects in time and space. Distant objects (apples in a faraway country) will behave like the ones one could actually observe. The same is true for objects separated in time. An observation made today (all other things being equal) can be repeated tomorrow, and the day after tomorrow and so on. But can the premise that unobserved objects resemble observed objects in all relevant aspects lead to justified true belief? The 18th-century Scottish philosopher David Hume, who was first to address this problem, suggested the addition of this general premise to all inductively derived knowledge: He calls the underlying principle allowing the generalization of a proposition made from the observation of a subset of objects or cases the *uniformity of nature* premise. The uniformity of nature assumption is no insurmountable rational justification for induction. As stated earlier, knowledge can be defined as *justified true belief*. By invoking the uniformity-of-nature assumption, inductive reasoning cannot be justified in a strictly logical sense. Although we might have good reasons to assume that the "next" observable object (e.g., another apple about to fall from a tree) will behave in a way already observed, this claim is only tenable under the assumption that nature is nonrandom and sudden changes in the behavior of observable objects do not occur.

Hume's problem cannot be solved in a strictly logical sense. Instead of proving the uniformity-of-nature assumption, scientists usually try to pile up empirical evidence to enhance the plausibility of the belief. Hume pointed out that this strategy itself entails an inductive argument. Any attempt to increase the plausibility of the uniformity-of-nature assumption by arguing that observed evidence is a reliable guide to what can be expected in the future simply repeats the uniformity-of-nature argument.

From an inductive perspective, the observation of some falling apples never allows us to formulate general laws of gravity (in a strict logical sense) although the uniformity-of-nature assumption gives us good reasons to expect the next apple to fall.

Deduction and the Covering Law of Explanation

In contrast to inductive reasoning, *deductive explanation* rests on the assumption that conclusions regarding specific cases can be derived from general premises and laws.

Among the most influential claims about deductive scientific explanation is the so-called *covering law of explanation* (also referred to as *deductive-nomological* or D-N model) introduced by Carl Hempel and Paul Oppenheim in 1948. In order to be *sound*, any scientific explanation must satisfy logical as well as empirical conditions of adequacy. Most important, the explanandum must be a logical consequence of the explanans; that is, it must be logically deducible. Therefore the explanans must contain general laws and empirical (testable) content. Moreover, to satisfy the conditions of empirical adequacy, the laws and statements entailed in the explanans must be “true.” That means that all the relevant available evidence supports them. The formal structure of the deductive-nomological model resembles the Aristotelian syllogism:

The explanans consists of one or more scientific laws (L) and a set of specific conditions (C):

$$\begin{array}{l} L_1, \dots, L_n \\ C_1, \dots, C_n \end{array}$$

The explanandum (E) must be logically deducible from the explanans. The model rests on the operation of general laws and empirical particular facts. Scientific explanation in this model means to demonstrate that a particular phenomenon can be subsumed under a covering law. The deductive-nomological model of explanation demands universality in a strict sense. Explanation must cover *all* cases with specific conditions and aims at deriving laws and general theoretical principles. If the predicted connection between the explanans and the explanandum fails to materialize in *some* cases, according to the deductive-nomological model, we cannot speak of a scientific explanation of the phenomenon in question at all. But is such a model sufficient for all kinds of circumstances?

Medical doctors are accustomed, for example, to attach a mortality rate to a particular disease.

Statistically, a certain percentage of patients who are given a specific medical treatment will still succumb to a disease. However widespread in everyday scientific practice, examples like this beg the question whether scientific explanation must establish universal, logical connections (Y is always a consequence of X) or if probabilistic statements (Y is the likely consequence of X) are sufficient.

Probabilistic Explanation

Instead of establishing universal lawlike connections, probabilistic explanations state that if certain specified conditions are fulfilled, the probability of a particular occurrence can be measured. In this model, the explanans does not logically/deductively imply the explanandum. In other words, the truth of the explanans (e.g., the correct diagnosis of a specific disease) does not make the truth of the explanandum certain (e.g., the consequences may vary between individual cases/patients). The probabilistic-statistical model leaves room for doubt. Instead, depending on observation and statistical calculation, the explanans confers a certain degree of likelihood on the explanandum. Similar to the deductive-nomological model, probabilistic approaches presuppose some general laws (in this sense they are also nomological); but dissimilar to D-N models of explanation, laws are treated as statistical connections rather than universal, logically necessary regularities.

Induction and Probability

Probabilistic explanation is nomological because it presupposes the existence of general laws. In contrast to the D-N model, these laws are of statistical rather than universal form. As a result, the explanatory arguments, which can be derived from statistical laws, are inductive in character. Whereas in the deductive perspective, a particular question is (or is not) covered by a universal law, inductive explanation leaves room for varying numerical degrees of empirical confirmation. Actually, probabilistic arguments are used to justify inductive reasoning, because the reliability of inductive reasoning increases with the statistical probability of a particular observation. While

the observation of a black swan (as in the example above) would, from the perspective of deduction, falsify any inductively derived claim about the color of these kinds of birds, it can still be reasonable to adhere to the claim if statistical calculations confirm its reliability. If nine of 10 swans actually are white, it is reasonable to assume that (the vast majority of) swans are white, or to speak of swans as white birds. Additionally, probabilistic explanation underlines the practical importance of inductive reasoning. Most often science aims not only at developing justified true beliefs but also at practical knowledge. If, for instance, a particular treatment proves (statistically) effective in 80% of cases of a disease, it would be highly impractical to conclude that there is no treatment available at all, although its appropriateness had been disproved in two of 10 cases.

Explanation and Prediction

Deductive and inductive reasoning aim primarily at scientific explanation. In both perspectives, observation must come first. Prediction, in contrast, aims at making scientifically sound statements of yet-unobserved phenomena. However, giving explanations and making predictions respectively should not be conceived as rival perspectives only. In an instrumental understanding of scientific theorizing, the ability to predict specific outcomes can—even more than its ability to deliver explanations—be the touchstone of the quality of a theory.

In general, deduction employs a logical dualism in regard to prediction: If a “new” case is covered by a particular law, its outcomes can easily be predicted; if it is not, then nothing at all can be said about the outcomes. Inductive reasoning offers good reasons (given the uniformity-of-nature assumption) to expect a specific outcome, in case its likelihood can be derived from a statistical law. Deductive reasoning predicts the same outcomes as have been confirmed for a class of phenomena/cases $C_{1...n}$ covered by a (set of) universal law(s). Inductive reasoning derives the likelihood L for a new Case C_{n+1} from the statistical connection already observed in a number of cases $C_{1...n}$. The limits of scientific prediction, which also separate scientific prediction from

prophecy, are defined by the availability of existing observation.

In practice, scientific predictions are of paramount importance while at the same time posing a major problem for the philosophy of science. The main problem is that predictions, in particular with regard to issues such as global climate change or international health risks, are based on models. Despite the complexity of contemporary modeling exercises, which allow the inclusion of thousands of variables and interactions, such models present simplified versions of reality and often involve deliberate omissions and iterations of the underlying set of variables to test for variations that affect the probability of outcomes.

Scientific Progress, Revolution, Evolution, Programs, and Traditions

Science is not static. Moreover, scientific knowledge must not be treated as a treasure that is constantly increased. Also, the process of scientific knowledge production is not a teleological venture, continuously diminishing the amount of ignorance until everything shall be known. Hypotheses can be verified *or* falsified, and at times theories have to be abandoned or reformulated. The history of science gives plenty of examples of change, even dramatic change: The Copernican turn in the mid-16th century symbolizes such a paradigm shift, replacing the Ptolemaic model that put the Earth at the center of the universe, with the new, heliocentric worldview still in place today. Discovery of the natural element of oxygen by Antoine-Laurent Lavoisier in the second half of the 18th century provided a better explanation for the process of combustion, rendering the then predominant “phlogiston theory” obsolete and led the way to modern chemical science. Albert Einstein’s theories of special and general relativity, developed in the early 20th century, radically changed modern physics. When scientific knowledge is subject to change, sometimes fundamental change, how does science then proceed? An influential conceptualization was developed by the American philosopher Thomas Kuhn. In his view, scientific progress is characterized by periods of stability with an accepted set of established theories and scientific practices, and periods of

scientific “revolutions,” in which time-honored concepts and ways of approaching the world are abandoned and eventually replaced by (radically) new ideas, theories, and practices. During periods of stability, a scientific discipline can rely on an established set of theories, ideas, and methods. Scientific inquiry during such periods of stability is referred to as “normal science.” Central to this concept of normal science is the existence of a *paradigm*. A paradigm consists of some fundamental theoretical assumptions and concepts commonly shared by the members of a particular discipline, together with some successful applications of the theoretical concepts. In order to become paradigmatic and transmitted in the textbooks of a discipline, a theoretical concept must be evidentially suitable to solve some empirical problems within the respective field of research. A paradigm normalizes a scientific discipline in a way that it provides a firm theoretical foundation and simultaneously represents a unifying outlook of how further research should be conducted. Day-to-day activity in the modus of normal sciences consists of the explanation of particular problems by employing the agreed-upon theories and methods covered by the ruling paradigm. Consequentially, Kuhn invokes the metaphor of the revolution to describe the moment when the unifying consensus is challenged and the ruling paradigm is tackled. Much as in the case of political revolution, the first question that must be answered refers to the timing of a revolution. In Kuhn’s view, normal science is engaged with trying to solve minor problems without introducing substantial change to the paradigm. Over time, anomalies may occur; that is, problems apparently inexplicable by the ruling paradigm. A sense of crisis may emerge when the number of anomalies increases and/or the problems of explaining them are directly linked to core features of the assumptions and concepts commonly shared. Under these circumstances, confidence in the old paradigm is weakened. This process gives way to the emergence of radical new approaches. In Kuhn’s words, science enters a *revolutionary phase*. During this time of disorder, various competing concepts emerge, while the scientific community, now divided into factions, is trying to regroup and “normalize” their research activity. Eventually the

best available concept will win over the majority of the representatives of a discipline and become the new paradigm, signaling the beginning of a new stable period of normal science. In sum, scientific progress in this model is characterized by lengthy periods of normal sciences, punctuated by occasional revolutionary phases.

The theory of scientific revolution was widely acclaimed for its ability to explain moments of paradigmatic shift in science, as in the preceding examples. At the same time, alternative theories of scientific progress were developed to cope with the apparent shortcomings of Kuhn’s approach. While the predominance of a single paradigm may be true in the natural sciences, other scientific disciplines (e.g., the social sciences) can be labeled as multi-paradigmatic. Also, there are examples of significant scientific progress without paradigmatic crises. Progress in particle physics in the last five decades or so, for example, has led to the refinement of the “particle zoo” metaphor and the development of a standard model without sparking paradigmatic shift. In other disciplines, particularly in most social sciences, ruling paradigms hardly ever exist. In those fields of inquiry, competing (or complementary) theories often coexist. From the perspective of Kuhn’s approach, they have to be labeled as “constantly revolutionary sciences” (which would miss much of the day-to-day activities of researchers in these disciplines) or deny their status as scientific disciplines altogether. In order to explain the continuous and incremental growth of scientific knowledge, which can be found in natural and social sciences alike, Popper suggested a model of scientific *evolution* rather than revolution. In line with his fundamental principle of falsification (discussed earlier in this entry), his evolutionary approach of scientific developments starts with a particular problem (P_1). The questions linked to the problem give rise to a tentative theory (TT), which is tested and modified by what he called error elimination (EE). The process of error elimination is more than the optimization of a tentative theory but may in turn lead to the appearance of new, unexpected problems (P_2). Here the process starts all over again. The evolution of scientific knowledge can be expressed as follows:

$$P_1 \rightarrow TT_1 \rightarrow EE_1 \rightarrow P_2 \dots$$

The discovery of P_2 itself indicates an increase of knowledge. In contrast to Kuhn, Popper's model does not distinguish between distinct periods of scientific process but assumes an incremental change. Another influential model for explaining scientific progress was developed by the Hungarian-born mathematician and philosopher of science Imre Lakatos. His approach is situated between the poles of Kuhn's revolutionary and Popper's evolutionary model. For Lakatos, science is dominated neither by monolithic paradigms nor by isolated, tentative theories. Instead science is organized around *research programs*. A research program is a general/super-theory or a sequence of theories. They consist of three elements:

1. A set of *fundamental assumptions*. These form the hard core, or *negative heuristics*, of a research program, for they cannot be abandoned without rejecting the research program altogether.
2. *Positive heuristics*. These consist of a set of beliefs linked to the core assumptions of a research program but also include suggestions or hints for how to deal with potentially refuting observational evidence. In Lakatos's words, they form a "protective belt" surrounding the hard core of the research program. The hard core of the "Copernican research program" mentioned earlier is the assumption that Earth and the other planets orbit a stationary sun, forming a solar system. However, to explain and calculate the movements of the planets, supplementary assumptions (e.g., the introduction of epicycles) are necessary. Inadequacies between the assumptions articulated in the research program and actual observation can be attributed to the positive heuristics.
3. The third element of a research program is a *series of theories* ($T_1 \dots T_n$). These theories are its specific instantiations and are developed by adding auxiliary clauses to previous theories. In contrast to Popper's paradigm, Lakatos argues in favor of a more incremental form of scientific change. Research programs can be either progressive or degenerative. If the addition of auxiliary hypotheses helps to enhance the explanatory or predictive power of a research program, it is *progressive*. The hard core of its

fundamental assumptions remains well protected by its protective belt, which allows the formulation of credible auxiliary hypotheses. When auxiliary hypotheses must be added to conceal more fundamental flaws or simply to provide some answers when faced with troublesome new evidence, Lakatos speaks of *degenerative* research programs. Its incapability to provide "good" explanations on the basis of its positive heuristics may eventually lead to the erosion of some core assumptions, which can mean the "fading out" of a particular research program. In sum, Lakatos argues for an evolutionary model of scientific progress. In his view, change in research programs takes place at their respective peripheries. In the process of scientific progress, auxiliary hypotheses can lead to an overall degeneration of the program, eventually leading to the weakening of its positive heuristics. Parallel to Kuhn's conception of research paradigms, the "hard core" of a research program remains stable until the program as a whole fades out and is replaced by more productive, progressive ones.

It is against this static view of the positive heuristics of research programs, but also because of Lakatos's (as well as Kuhn's) emphasis on empirical problems as the driving force behind scientific growth, that the American-born philosopher Larry Laudan developed his own concept of research traditions. While he agrees that ultimate aim of science is the development of theories with high problem-solving capacities, his model differs from those of his predecessors in some major ways: He is fundamentally skeptical toward the belief that theories can deliver accurate insights about the way the world actually is. He attempts to circumvent the problem of the truthfulness of a theory by focusing on its problem-solving capacity. According to Laudan, unlike Kuhn and Lakatos, the quality of a theory is not solely dependent on its ability to solve empirical problems. He suggests a taxonomy that distinguishes between empirical problems on the one hand and conceptual problems on the other. Empirical problems can be further differentiated as solved, unsolved, and anomalous problems. By solved problems he refers to empirical questions adequately solved by an existing theory. Unsolved problems lack resolution by any

existing theory, while anomalous problems have been solved by a particular theory but cannot be explained by one or more rival theories. As part of the conventional wisdom, unsolved problems are regarded as the major stimulus for scientific progress. Finding a solution where none has existed drives the development of new theoretical approaches. Conceptualizing scientific progress as filling the white spaces on the map of knowledge by solving more and more problems is, in Laudan's view, too short-sighted. Likewise, changing the paradigm or adding auxiliary hypotheses to deal with unexpected (unsolved) problems falls short of explaining the nature of scientific growth. Focusing on the problem-solving capacity of theories with regard to empirical problems neglects the second category, which entails conceptual problems. Conceptual problems arise when a theory (t_1) is in conflict with another theory (t_2), which is believed to be rational and well-founded. The simplest form of conflict between theoretical concepts is their logical inconsistency or incompatibility. If t_1 entails the negation of (crucial parts of) t_2 it speaks of inconsistency, whereas in the case that t_1 employs a model of explanation that entails nothing about t_2 , the two concepts are regarded as incompatible. Apart from such intrascientific difficulties, conceptual problems can result from normative difficulties (including methodological rules as norms of good scientific behavior) or contradictions to nonscientific beliefs or metaphysical assumptions (worldviews). Laudan's term of research traditions aims at providing a criterion for theory selection given the conceptual and empirical problems outlined above. He describes research traditions as guidelines for the development of specific theories, a specific set of beliefs about the kinds of entities and processes that constitute the domain of inquiry. Additionally, research traditions entail one or more epistemic and methodological norms, which define how a specific field of inquiry must be investigated and which methods are considered legitimate.

Scientific progress, especially the deviation from existing theories and models, is, in Laudan's view, not exclusively triggered by the shortcomings of providing answers to empirical questions. Instead, even empirically well-supported theories can be abandoned in favor of less supported approaches, provided that these are equipped to

resolve conceptual difficulties confronting their predecessors.

The concept of research tradition shares a common assumption with the models of research programs and scientific revolution, namely the conviction that at any given time, a (more or less) defined set of theoretical and epistemic assumptions exists and that "good reasons" (unsolved empirical and/or conceptual problems) are necessary for abandoning theoretical concepts.

A radically contradicting view was introduced by the philosopher of science Paul Feyerabend (1924–1994). In his controversial book, *Against Method*, he argued for epistemological anarchy, a position famously culminating in his principle of "anything goes." From this perspective, scientific progress cannot be produced by methodological monism. He intended to show that compliance with the principles of methodological monism (interdiction of ad hoc hypotheses, increase of empirical content, theoretical consistency) and in particular the principle of falsification imply the impossibility of scientific progress. The principle of "anything goes" does not call for the inclusion of antiscientific methods in scientific reasoning and is no methodological principle itself, but argues against methodological rigidity and a close relation between theoretical and methodological purity and scientific progress. From this perspective, the advancement of scientific knowledge depends on a variety of factors including chance and the breaking of scientific rules, as well as non-scientific aspects (such as aesthetic criteria and social factors).

Natural and Social Sciences

So far, no distinction has been made between natural and social sciences. According to the thesis of the unity of science, all scientific disciplines are part of the same holistic endeavor to gather scientific knowledge: Common scientific laws form the basis of any scientific inquiry. However, the nature of the field of inquiry of social sciences poses some important problems for that thesis and the philosophy of sciences in general.

A longstanding question is therefore if, and to what degree, social sciences resemble natural sciences. The 19th-century French philosopher Auguste Comte imagined the emergence of

“positive” social sciences, which should develop in the image of the natural sciences. Like any other discipline, the social sciences in his view have to pass through three distinct stages: a *theological* phase, in which explanation of natural occurrences (e.g., weather events) and social occurrences (e.g., victories in battles, but also one’s personal fortune) are explained by the interference of personal deities. The initial stage is followed by a *metaphysical* phase, in which the personal and indirect influence of gods is replaced by abstract, impersonal concepts, which can be rationally analyzed. The final phase is called the *positive* stage, which is dominated by scientific explanation based on observation. Comte’s conviction that social sciences can become a kind of social physics occupied with uncovering “social laws” is disputed by many of today’s philosophers of science. Social sciences are seen as fundamentally different from natural sciences, mainly for two reasons: (1) the difference of explanation and understanding as major aims of scientific inquiry, and (2) the constructed nature of the social world and the responsiveness of societies and individual social actors to scientific findings.

One cardinal difference between natural and social sciences is that of their respective aims, stemming from the problem of classification and the historicity of cases. While both the French Revolution of 1789 and the Arab Spring since 2010 might be labeled as “revolutions,” the semantic similarities must not conceal the fundamental differences between the two cases. It is therefore questionable whether some “general laws of revolution” (explaining the causes and courses of any political upheaval) can be formulated at all or whether each case has to be “explained” by drawing on respective historical and cultural peculiarities. While the identification of laws aims at explaining underlying principles and mechanisms, social sciences, in the words of the classical German Sociologist Max Weber, seek to understand the characteristics of particular cases.

Moreover, unlike in natural sciences, social facts are constituted by societies themselves. Social laws are created (constructed) by social actors and therefore can (deliberately) be changed. Additionally, social actors can respond to the

findings and predictions of the social sciences. As the American sociologist Robert Merton stated, in the social world, scientific description can even prescribe the developments it claims to have discovered (self-fulfilling prophecy). Both of these problems limit the use of nomological-deductive reasoning in social sciences, for neither can universal laws simply be discovered nor is it always possible to explain the interplay between structure and agency by referring to universal social laws. Statistical probability and inductive reasoning, in general more important to social sciences, also has to be treated carefully. The statistical connection of a person’s educational background and her ability to reach a privileged position, for example, can change when some structural features of the society in question are altered (e.g., special programs for female students) but also when the assessment of the available career opportunities change (e.g., the role of women in general). Philosophy of social sciences today distinguishes between naturalist positions stating that the methods of social sciences should correspond to those of the natural sciences, and anti-naturalist positions, which emphasize the differences between understanding (predominant in the social sciences) and explanation (the main goal in natural sciences), which requires the inclusion of non-naturalistic methods (e.g., hermeneutics) of inquiry.

Value-Free Science and Scientific Authority

Feyerabend’s aforementioned critique of methodological monism in science is echoed by contemporary philosophers and sociologist of science. Since the 1970, postmodernists such as Jean-François Lyotard, proponents of the “strong program” in the sociology of knowledge and science and technology studies (STS), have emphasized the social role of science. With the increasing role of science-driven technological change and the paramount importance of scientific knowledge in modern societies, philosophy of science emphasizes the role of values *in* and *for* the authority of science. The role of norms and values in science has gained additional attention since scientific modeling and forecasting have

become ever more important. As mentioned earlier in this entry, understanding complex and consequential developments such as climate change involve the use of models. However, the need to decide which variables should be included invites the criticism that scientists “tailor” their model assumptions to derive predictions that correspond with normative, economic, or political preferences.

In addition to “cognitive” or “epistemic” values such as accuracy, simplicity, or coherence with accepted theories, proponents of STS and the strong program emphasize the role of contextual values in science. In contrast to epistemic values—which directly refer to, and can be derived from, the immanent logic of research—contextual values include political, cultural, and personal values and can thus compromise scientific objectivity and in consequence scientific authority. For if scientific knowledge is reflecting the cultural or personal norms of the researchers, the objectivity of scientific knowledge can be called into question. Science should then be regarded as a mere social practice that reflects the preferences and prejudices of the involved scientists.

Disputes about scientific objectivity and the role of contextual values for the production of scientific knowledge highlight new cleavages within the philosophy of science, which is increasingly concerned with understanding the political and social impacts of, and on, science. At the same time, the social role of philosophy of science itself changes: No longer limited to understanding the intricacy and inner workings of scientific research, contemporary philosophers of science increasingly reflect upon the integrity, efficacy, and legitimacy of science in modern societies.

Alexander Ruser

See also Knowledge; Prediction; Pseudoscience; Reasoning

Further Readings

- Boyd, R., Gasper, P., & Trout, J. D. (1991). *The philosophy of science*. MIT Press.
- Curd, M., & Cover, J. A. (1998). *Philosophy of science: The central issues*. W.W. Norton.

- Douglas, H. (2010). Engagement for progress: Applied philosophy of science in context, *Synthese*, 177, 317–335. <https://doi.org/10.1007/s11229-010-9787-2>
- Elliot, K. (2011). Direct and indirect roles for values in science. *Philosophy of Science*, 78(2), 303–324.
- Feyerabend, P. K. (2010). *Against method* (4th ed.). Verso.
- Hempel, C. G., & Oppenheim, P. (1948). Studies in the logic of explanation. *Philosophy of Science*, 15(2), 135–175.
- Kuhn, T. (1962). *The structure of scientific revolutions*. University of Chicago Press
- Laudan, L. (1977). *Progress and its problems. Towards a theory of scientific growth*. University of California Press
- Popper, K. (1959). *The logic of scientific discovery*. Basic Books.

PHYSICAL THEORY

Physical theory is concerned with the fundamental constituents of the universe, the forces they exert on each other, and the results produced by said forces. It is closely related to the other natural sciences and in many ways encompasses them. Chemistry, for example, deals with the interaction of atoms to form molecules, geology is largely a study of the physics of the world, and astronomy deals with the physics of outer space.

A physical theory is a model of physical events and cannot be proven from basic axioms. It differs from a mathematical theorem since physical theories model reality and are a statement of what has been observed. They can also provide predictions about new observations. There is therefore more to a physical theory than application or even invention since physical theories are involved in concept formulation. Very often, it is a mathematician who provides the clue to a physical theory—for example, Bernhard Riemann’s notion that space might be curved. This entry presents an overview of the foundations of physical theory and a history of its advances from origins in ancient Greece through the medieval and Renaissance periods to the present day. The entry concludes with a description of the problems and challenges facing contemporary physical theorists and researchers in their efforts to produce a

theory that unites the four known fundamental forces into a unified field theory.

Foundations

The advancement of science depends on the interaction between physical theories and experimental physics. Sometimes, physical theory develops mostly along the standards of mathematical rigor giving less weight to experiments and observations. For example, when developing special relativity, Albert Einstein concerned himself more with the Lorenz transformation than with the Michelson–Morley experiment that attempted to prove the theory. Another time, however, he paid much more attention to experiments, for example, when formulating the photoelectric effect whereby the existing experimental results required a theoretical formulation.

A physical theory is involved in the relationships between different kinds of measurable quantities. This is exemplified by the Greek physicist and mathematician Archimedes (c. 287–212 B.C.E.), who came to realize that the volume of an irregularly shaped object could be determined by immersing it in a liquid, and that a ship is able to float due to the displacement of its weight of water. The Ionian Greek philosopher Pythagoras (c. 570–495 B.C.E.) identified the relationship between a length of vibrating string and its effect on musical tone as well as how to work out the length of a rectangle's diagonal. Thermodynamics is a branch of physics that employs four rules. The second of these involves a quantity termed *entropy*, which is usually symbolized by the letter *S*, and is a measure of the number of specific ways in which a thermodynamic system can be arranged. The modern microscopic interpretation of entropy in statistical mechanics involves entropy providing the amount of additional information required to specify the exact physical state of a system given its thermodynamic specification. Entropy is often regarded as an indication of the randomness or disorder of a system.

Theoretical physics can take a number of approaches. Those working in theoretical particle physics, for example, might use empirical formulas to agree with experimental results and can do prior to achieving an understanding of the physics

behind them. *Modelers*, those who build models, may try to model speculative theories so they possess certain desirable features rather than basing them on experimental data, or they may apply the techniques of mathematical modeling to physics problems. Some theoretical physicists try to create approximate theories, known as *effective theories*, seeing fully developed theories as either unsolvable or overly complicated. Effective theories are essentially approximations to underlying physical theories. These include quantum field theory or a statistical mechanics model. These types of physical theories include degrees of freedom to describe physical phenomena that occur at selected length or energy scales while ignoring degrees of freedom at shorter distances (or higher energies). Effective theories serve to simplify calculations and enable physicists to address the effects of radiation and dissipation.

Others working in the area of physical theory may attempt to formalize, unify, or reinterpret existing theories or create completely new ones. Theoretical advances in physics often consist in putting aside older paradigms. Pure mathematics can sometimes provide clues about how a physical system can be modeled while others require computational investigation.

Physical theories become accepted if they make correct predictions and no (or few) incorrect predictions. *Occam's razor* is a principle that is often used as an injunction not to make more assumptions than you actually need. It's named after 13th-century philosopher William of Ockham, which states that the simpler of two theories that describe the same subject just as well is to be preferred. This is only true if two rival theories predict identical results but infer the same results. In science, however, it's more often the case that theories are distinguished less by making fewer assumptions *per se*—but rather by making different assumptions.

When attempting to develop a testable physical theory, scientists should follow the concepts defined in the scientific method. The scientific method is a set of techniques used by the scientific community to investigate natural phenomena by providing an objective framework in which to make scientific inquiry and analyze the data to reach a conclusion about that data. Modern

theoretical physics tries to bring theories together and to explain phenomena in even further attempts to understand the Universe, both from a cosmological point of view and on an elementary particle scale. When experimentation is not possible, theoretical theories can still attempt to further knowledge via the use of mathematical models.

Famous theoretical physicists include Isaac Newton, Albert Einstein (1979–1955), Max Born (1882–1970), Max Planck (1858–1947), Paul Dirac (1902–1984), Richard Feynman (1918–1988), Abdus Salam (1926–1996), Enrico Fermi (1901–1954), and Stephen Hawking (1942–2018).

Categories of Physical Theories

Physical theories can be grouped into three categories: *mainstream*, *proposed*, and *fringe* theories.

Mainstream Theories

Mainstream theories are sometimes referred to as *central theories*. They are the body of knowledge of both factual and scientific views and possess the scientific qualities of test repeatability; consistency with existing, well-established scientific theories; and experimentation. There are some mainstream theories that are generally accepted based on their effects explaining a large variety of data, while the detection and explanation and possible composition are open to debate. Examples of mainstream theories are *cosmology*, *dynamics*, *general relativity*, *quantum mechanics*, *string theory*, *thermodynamics*, and *particle cosmology*.

Proposed Theories

Proposed theories of physics are generally quite new theories that deal with the study of physics using scientific approaches, a means for determining the validity of models and new types of reasoning used to arrive a theory. However, there are some proposed theories that have been in the mix for decades and have so far eluded a way of testing and discovery. Proposed theories often include fringe theories that are undergoing the process of establishment and as yet are waiting to gain wider acceptance. Proposed theories are usually untested theories. Some examples

include *dark energy*, *Einstein–Rosen bridge*, *M-theory*, and *emergence*.

Fringe Theories

Fringe theories include new areas of scientific endeavor, those that are in the process of becoming established as well as some proposed theories. These can include speculative sciences in the fields of physics and physical theories, submitted in accordance with known evidence where a body of associated predictions has been put forward according to those theories. Some fringe theories subsequently become part of the widely accepted theories of physics. Others go on to be disproven. Certain fringe theories are forms of protoscience, while others are a form of pseudoscience. Disproving the original theory may lead to a reformulation of the theory. Examples of fringe theories include the dynamic theory of gravity, grand unification theory, luminiferous aether, steady state theory, and the theory of everything.

History of Physical Theory

The Babylonians, ancient Egyptians, and early Mesoamericans were early observers of planetary activity but failed to find an underlying system for planetary motion. The earliest known form of physical theory began more than 2,300 years ago in pre-Socratic Greece and was continued by Plato and Aristotle, some of whose ideas remained the mainstay for a millennium. The downside to Greek physics was that it discouraged experimentation, something which unfortunately held back science as a whole during that time. In the medieval era, when the first universities were being built, the acknowledged disciplines were mathematics, theology, medicine, and law. As the concepts of energy, space, matter, and causality developed, physics began to acquire the form of discipline we have today. During the Renaissance the modern-day concept of experimental science as a counterpoint to theory first emerged with Francis Bacon (1561–1626), who advocated the experimental method as the true foundation of scientific knowledge. He also worked on astronomy, optics, chemistry, and machine design. The Copernican Revolution was a landmark in the

history of physical theory when the mathematician and astronomer Nicolaus Copernicus (1473–1543) proposed a heliocentric system, in which the planets move around the sun. This caused a paradigm shift in astronomy that heralded the modern era in physical theory. This was soon followed by Johannes Kepler's (1571–1630) theory of planetary orbits based on Tycho Brahe's (1546–1601) meticulous calculations.

Galileo Galilei (1564–1642) was instrumental in the push toward modern concepts of explanation. As a physicist he was theoretically and experimentally gifted. He was able to confirm the heliocentric model by observing the phases of the planet Venus. He also discovered surface regularities on the moon using his newly developed telescope. Galileo was not just an astronomer, he also performed experiments that demonstrated that bodies of different weights fall at the same rate, overturning Aristotle's false dictums, and that their speed increases at a uniform rate with the time of the fall. His work foreshadowed the work of 17th-century physicist and mathematician Isaac Newton.

Newton, Classical Mechanics, and Fundamental Forces

At around the age of 23, Newton propounded the principles of mechanics, proposed a theory for the propagation of light, formulated the law of universal gravitation, and invented differential and integral calculus. His contributions to physical theory were enormous. He succeeded in illustrating that Kepler's laws of planetary motion and Galileo's theories involving falling bodies were in line with his own laws of motion and gravitation. He also predicted the appearance of comets and explained the effects of the moon on the tides explaining the precession of the equinoxes. Physics subsequently owed much to Newton, and in particular to the second law of motion which states that the force needed to accelerate an object is proportional to its mass times the acceleration. If the force and initial position and speed are known, subsequent positions and velocities can be worked out. The force may vary with time or position, in which case Newton's calculus needs to be applied. Newton's simple law was even more far-reaching in that it stated that each body possesses

an inherent property, its inertial mass, and it is this that influences its motion. The greater the mass the slower the change of velocity when impressed by a given force. This law has a practical aspect today, as long as the body discussed is not overly small, not overly big, and not moving very quickly. Newton's third law is expressed simply as "for every action there is an equal and opposite reaction"; that is, all forces between particles arise in oppositely directed pairs.

Newton's force of gravity was his chief contribution to the description of natural forces. The law of gravitation recognizes that all material particles and the bodies that are composed of them have a property termed *gravitational mass*. This causes any two particles to exert attractive forces on each other along the line joining them, in direct proportion to the product of the masses, and inversely proportional to the square of the distance between them. This law governs the motion of the planets around the sun, and the earth's own gravitational field. It is also strongly suspected to be involved in the possible gravitational collapse seen in the final life cycle of stars.

Nowadays, we know that there are three other fundamental forces in addition to gravity that are active in the universe:

Electromagnetism, which is manifested in the interaction between particles with an electrical charge. Charged particles at rest interact via electrostatic forces. In motion, they interact through both electrical and magnetic forces. For many years, electromagnetic forces were thought to be different from one another but in 1864 physicist James Maxwell unified them under Maxwell's Equations. In the mid-20th century, quantum electrodynamics combined electromagnetism with quantum physics. Electromagnetism is very evident in the world since it can affect things that are at a distance with a reasonable amount of force.

Weak interaction is a powerful force that acts on the scale of the atomic nucleus and causes such phenomena as beta decay. Weak interaction has been joined with electromagnetism in a sole interaction known as the *electroweak interaction*.

Strong interaction is the strongest of the forces that keeps nucleons (protons and neutrons) bound together. For example, it has the strength to bind

two protons together in a helium atom, even though the positive charges in the atom cause them to repulse each other. The strong interaction enables particles termed gluons to bind quarks together to create the initial nucleus. Gluons can also interact with other gluons giving the strong interaction a theoretically infinite distance.

Many physicists believe that all four fundamental forces belong to a single underlying force, which has yet to be discovered. In the same way as electricity, magnetism, and the weak force were conjoined in the electroweak interaction, research work continues to try to unify the fundamental forces. Current quantum mechanical thought is that the forces' particles do not directly interact but rather produce virtual particles that mediate the actual interactions. All forces apart from gravity have been incorporated in this standard model of interaction.

René Descartes' (1596–1650) analytic geometry and mechanics became incorporated into Isaac Newton's calculus and mechanics. Descartes provided the first truly modern formulation of the laws of nature and, in addition, constructed an influential theory of planetary motion. He rejected the use of substantial forms, instead favoring a mechanical approach to the explanation of natural phenomena, which entailed the interaction of small unobservable *corpuscles* of matter. One of his chief accomplishments was to discover the laws of refraction, commonly termed *Snell's law*, which pronounces that as light passes from one medium to another the sine of the angle of incidence maintains a constant ratio to the sine of the angle of refraction. His ambitious text entitled *The World* contained the blueprints for a mechanical/geometric physics and set out his vortex theory of planetary motion. In 1644, he published an exhaustive investigation into physics in the *Principles of Philosophy*, which also detailed the metaphysical underpinnings of his physical system.

Modern Physics

Joseph-Louis Lagrange, Leonhard Euler, and William Rowan Hamilton were responsible for extending the theory of classical mechanics, having picked up the thread interweaving mathematics

and physics started by Pythagoras more than two millennia before. The Lagrangian mechanics is a reformulation of classical mechanics, using the principle of least action. Joseph-Louis-Lagrange introduced this system in the late 18th century. The Lagrangian is a mathematical function summarizing the dynamics of the system. In a simple mechanical system, it is the value given by the kinetic energy of the particle less the potential energy of the particle; however, the principle can be extended to complex systems. This reformulation was an alternative response to Cartesian coordinates including the *polar*, *spherical*, and *cylindrical* coordinates that were found to be unsuitable for Newtonian mechanics. The Lagrangian has been used in Newtonian physics as a way to find the acceleration of a particle in a gravitational field. It has also been used in Einstein field equations, something that led to its being applied in electromagnetism, to curved space-time, and when describing black holes.

Some of the best conceptual achievements of the 19th and 20th centuries include the consolidation of the idea of energy by the inclusion of heat, and subsequently electricity, light, and mass. The laws of thermodynamics, in particular the introduction of a singular concept of entropy, filled in large gaps in the attempt to explain “why things happen.”

The most revolutionary and pivotal theories of modern physics are the *theory of relativity* and *quantum physics*. Newton's mechanics were integral to special relativity, and his gravity was accorded classical mechanical explanation by general relativity. The area of quantum mechanics was fundamental to the ensuing understanding of *blackbody radiation* and of unusual instances in the specific heats of solids. Eventually, this work would lead to an understanding of the internal structures of molecules and atoms.

Quantum Physics

Quantum physics represents the sum total of our knowledge about how matter and energy interact in the universe. Quantum physics, also known as *quantum mechanics* or *quantum field theory*, is a molecular study of atomic, nuclear, even tinier microscopic levels. The term *quantum*

means “how much” and refers to the discrete units of matter and energy that are predicted by and observed in quantum physics. Space and time, which we usually think of as continuous, also have the minutest possible values. The birth of quantum physics is attributed to the German physicist Max Planck (1858–1947) in his paper on blackbody radiation, published in 1900. The system was further developed by Albert Einstein, Werner Heisenberg, Erwin Schrödinger, and Niels Bohr, among others. Einstein actually had serious theoretical issues with the field and attempted over a period of many years to disprove or revise it.

Quantum physics deserves especial attention because it’s the science behind how everything works, providing a description of the nature of particles and how they interact. Light waves perform like particles, and particles act in the same way as waves. Matter can apparently travel from one spot to another without moving through the intervening space, something known as quantum tunneling. Information moves vast distances instantaneously. A key concept is that of *quantum entanglement* whereby multiple particles associate so that measuring the quantum state of a particle places constraints on measuring the others. This was first demonstrated by the EPR Paradox, and while this concept started as a ‘thought experiment’ it has since been confirmed via experiments, tests known as Bell’s Theorem.

Thought experiments are situations created in the mind that ask questions such as “suppose you are in a certain situation, assuming such and such is true—what would happen?” Thought experiments are sometimes created to investigate phenomena that cannot readily be experienced in everyday life. Some famous examples of thought experiments include Schrödinger’s Cat, and as mentioned, the EPR thought experiment. Thought experiments usually lead to real experiments designed to verify the conclusions, and therefore the assumptions of them. The EPR thought experiment was devised by Einstein and colleagues Boris Podolsky and Nathan Rosen as a critique against the Copenhagen interpretation of quantum mechanics.

Quantum optics is another branch of quantum physics that is concerned with the behavior of light or photons. The behavior of photons has a

bearing on out-coming light. Lasers are one important application originating from quantum optics.

Quantum electrodynamics, developed in the mid-20th century by Richard Feynman, Julian Schwinger, and others, studies how electrons and photons interact; *unified field theory* attempts to bring together quantum physics and Einstein’s theory of general relativity by seeking to unify the fundamental forces of physics. Unified field theories include *quantum gravity*, *string theory*, *superstring theory*, *M-theory*, *grand unified theory*, *theory of everything*, *loop quantum gravity*, and *supersymmetry*.

All these massive achievements were dependent on theoretical physics as a moving force to both suggest experiments and consolidate results. Often, the ingenious application of mathematics was a prerequisite, and in the case of Descartes and Newton, the invention of new mathematical systems. Fourier’s studies on heat conduction led to a new branch of mathematics, the infinite orthogonal series. Heat conduction is one of the three modes of heat transfer. In 1822, French physicist Joseph Fourier set out a treatise that gave a complete theory and mathematical model of heat conduction. Heat conduction involves the transfer of heat from a warmer to a colder object by direct contact. He discovered that heat flux is in proportion to the size of a temperature gradient, and his equation became known as Fourier’s law.

A Theory of Everything

One goal of modern physics is to develop a comprehensive “theory of everything,” which reconciles gravity with the other three fundamental forces. If this could be developed, it would have to define the tentative boundary between Einstein’s theory of relativity and quantum theory, so this “theory of everything” is referred to as *quantum gravity*.

Quantum gravity is the overall term for physical theories that try to unify gravity with the other fundamental forces of physics, which are already united. It generally supposes the existence of a theoretical entity—a graviton or virtual particle that mediates the gravitational force. The

traditional model of quantum physics developed in the early 1970s propounds that the three other fundamental forces of physics are reconciled by *virtual bosons*. Bosons are types of particles that act according to the Bose–Einstein statistics. They possess a quantum spin and are sometimes known as *force particles* since they control the interaction of physical forces. The name *boson* originates with the Indian mathematician Satyendra Nath Bose (1894–1974), who worked with Einstein early in the 20th century.

Photons are elementary particles that are known as the *quantum* of electromagnetic food. Unlike other forms of matter, they cannot be subdivided, owing to their properties of waves—they can be refracted by a lens, interfering with other waves. Photons do not decay like other particles, and they carry no electrical charges. Photons mediate the electromagnetic force, while W and Z bosons mediate the weak nuclear force. *Gluons* (e.g., quarks) are force-mediating particles and exist in every atomic nucleus, holding it together. They mediate the strong nuclear force, which is stronger than electromagnetism and gravity by a long way. The *graviton*, being a theoretical particle with no mass and no charge, mediates the gravitational force. This theoretical particle is also believed to have spin two (spin being an intrinsic form of momentum) due to gravity's being described as having a certain kind of tensor field or mathematical framework.

Proving quantum gravity is experimentally problematic because the energy levels required to observe the postulates of the theory are unachievable in current laboratory experiments. In fact, quantum gravity has theoretical problems too. Gravitation is now explained by the theory of general relativity, which sets out different assumptions about the universe at a macroscopic level compared with the assumptions made by quantum mechanics on a macroscopic level. Within theorists' efforts to combine these, the problem of renormalization rears its head, whereby the sum of all the forces do not cancel out, resulting in an infinite value. Should quantum gravity actually exist, it is also possible that it will further our understanding of space and time, another reason why physicists continue to exert their efforts in pursuing this aim. Some

unified theories that are termed quantum gravity theories include supergravity, twistor theory, noncommutative geometry, and Euclidean quantum gravity.

Physical Theory Today

Theoretical physicist Lee Smolin suggests the five major problems in theoretical physics that are still giving physicists a perplexing time today. These are: 1. *The problem of quantum gravity*. Quantum gravity is the attempt in theoretical physics to create a theory that includes both general relativity and the standard model of particle physics. At present, these theories describe different scales of nature and attempts to explore areas where they overlap have not managed to yield any meaningful results, for example, the force of gravity (or curvature of space time) becoming infinite. This is a problem since physicists cannot visualize or want to visualize any infinities in nature.

2. *Foundational problems in quantum mechanics*. There still needs to be an understanding of the underlying physical mechanism in quantum mechanics. While there are numerous interpretations including the Copenhagen interpretation, the participatory anthropic principle, and Hugh Everett's many worlds principle, there is still disagreement about what causes the collapse of the quantum wave function. Most modern physicists working in quantum theory do not consider these questions of interpretation relevant. For many, the principle of de-coherence is the explanation, that is, interaction with the environment causes the quantum to collapse. Perhaps more significantly physicists can solve the equations, perform experiments, and practice physics without needing to resolve these types of fundamental questions.

3. *The unification of particles and forces*. Controversy still rages about whether or not the various particles and forces can be unified in a theory that can explain them all as the product of a single fundamental entity. As we have already seen, there are four fundamental forces of physics, but the standard model of particle physics includes just three (electromagnetism, strong, and weak nuclear forces). Gravity is not included in the standard model; attempting to create one theory that unites

the four forces into a unified field theory remains a goal in modern-day theoretical physics. Because the standard model of particle physics is a quantum field theory, any unification would therefore have to include gravity as a quantum field theory. This means that solving the third problem is connected with solving the first problem on this list. Additionally, the standard model of particle physics indicates a large number of particles, in fact 18 in total. Many physicists believe that a fundamental theory of nature should possess a method of uniting these particles so they can be described in more fundamental terms. String theory is perhaps one of the best defined of these approaches; it maintains that all particles are varying vibrational modes of filaments of energy or string.

4. *Problem of tuning.* A theoretical physics model is a framework that requires that certain parameters be set. In standard particle physics, the parameters consist of the 18 particles asserted by the theory, so the parameters are measurable by observation. Some physicists believe, however, that it should instead be the fundamental physical properties of the theory that determine the parameters, independent of measurement. This has motivated much enthusiasm for a unified field theory, even sparking Einstein's famous question "did God have any choice when he created the universe?" The question relates to whether the properties of the universe inherently provide the form for the universe, which begs the question whether these properties would work if the form were different. Many answer this by inclining toward the idea that there is potentially more than one universe that can be created along with a wide number of fundamental theories based on different physical parameters that could be the case.

5. *The problem of cosmological mysteries.* There are plenty of mysteries still remaining, but those that perplex physicists the most are the ones to do with *dark matter* and *dark energy*. This type of matter and energy can be detected by gravitational influences only and not by direct observation. Physicists still struggle to ascertain what they are, with some proposing alternative explanations for the gravitational influences, explanations that no longer require new forms of matter and energy.

Thora L. Fitzpatrick

See also Physics, 19th Century; Physics, 20th Century; Physics, Antiquity; Physics, Contemporary; Physics, Enlightenment; Physics, Gravitation Theory; Physics, Middle Ages; Physics, Quantum Theory; Physics, Renaissance

Further Readings

- Duhem, P., & Maurice, M. (1991). *The aim and structure of physical theory*. Princeton Science Library.
- Feynman, R., Leighton, R. B., & Sands, M. (1966). *The Feynman lectures on physics*. Addison-Wesley. <https://feynmanlectures.caltech.edu/>
- Garber, D. (1992). *Descartes' metaphysical physics*. University of Chicago Press.
- Hartmann, S. (2001). Effective field theories, reductionism and scientific explanation. *Studies in History and Philosophy of Modern Physics*, 32B, 267–304.
- Longair, M. S. (2003). *Theoretical concepts in physics: An alternative view of theoretical reasoning in physics*. Cambridge University Press.
- Planck, M. (1997). *Eight lectures on theoretical physics* (A. P. Wills, Trans.). Dover Books. (Original work published 1909).
- Smolin, L. (2006). *The trouble with physics*. Houghton Mifflin.

PHYSICS, ANTIQUITY

What we call physics today did not exist as a distinct unified discipline in the ancient cultures we know about. Our discipline is concerned with energy, with matter in motion and the forces that generate that motion, and with the fundamental constituents of the world. In the West, physics develops out of what was called *natural philosophy* until the late 19th century and focuses on the study of "the workings of nature." Its development into the discipline now called *physics* drew from the studies of optics, mechanics, geometry, and astronomy. These mathematical disciplines began in Western antiquity with Hellenistic writers such as Ptolemy and Archimedes and find some important precursors in Babylonian and Mesopotamian investigators. In addition to these roots in the west, important traditions for mathematizing nature are found in India and China.

Particularly important for the development of physics is the development of astronomical theories. This entry will focus on Western antiquity and its early theorizing about the workings of the natural world and then examine some of those elements from China and India.

Explanation in Early Physical Theory

Antiquity (the time preceding the Middle Ages) is generally the time we believe people began to diligently try to understand the behavior of matter. The mystery of the universe, its structure and origin, also came into serious question at this time. People wanted to understand the form of Earth and how the Sun and Moon behaved. Several theories were proposed during antiquity, and although most of them proved to be incorrect, some of those ideas hold true today. Most theories in antiquity were stated in philosophical terms, and they were typically not verified by testing. Telescopes and time-keeping equipment were not yet available.

Before antiquity, the nature of the Earth was typically explained by the actions of the gods. However, around the first millennium B.C.E., Greek philosophers began to argue that the world could be understood as a result of natural processes. They challenged typical ideas from mythology (such as the origin of the human species, though this was the origin of biology, not physics) and proposed regular underlying processes as themselves responsible for the things we see.

Ancient Greece

The Archaic period in Greece (occurring from 700 to 480 B.C.E.), the time of the pre-Socratic philosophers, marks the first known move toward a rational understanding of nature. Thales of Miletus (7th and 6th centuries B.C.E.), a philosopher known as the *Father of Science*, refused to accept mythological, supernatural, or religious explanations for natural phenomena and instead argued that every event had a natural cause. Thales also suggested in 580 B.C.E. that water is the basic element. He experimented with magnets and the attraction to rubbed amber and formulated the first cosmologies.

Thales's theories were not upheld for long, however, as Anaximander disputed his ideas and asserted that a substance called *apeiron* was the building block of all matter—not water.

Another standout philosopher was Heraclitus, who around 500 B.C.E. asserted that change was the only basic law governing the universe and that nothing ever remains the same. He was one of the first scholars to study the role of time in the universe, one of physics' most important concepts, even today.

The ancient Greek physicist Leucippus, who lived in the first half of the 5th century B.C.E., argued against divine intervention in the universe and argued that every natural event had a natural cause. Along with his student, Democritus, he was the first physicist to develop the theory of atomism, or the idea that everything is composed of indivisible elements called *atoms*.

Aristotle

During the classical period in Greece (6th, 5th, and 4th centuries B.C.E.) in Hellenistic times, physicists such as Aristotle (384–322 B.C.E.) advanced the study of natural philosophy. Aristotle, himself a student of Plato (ca. 429–347 B.C.E.), who advanced theories about geometrical objects as the constituents of matter, believed that the observation of physical phenomena could lead to the discovery of the natural laws surrounding them. In addition to his writings on poetry, metaphysics, logic, music, theater, politics, rhetoric, zoology, biology, ethics, and government, Aristotle wrote the first work that is referred to as *physics*, which is today known as *Aristotelian physics*. Aristotle explained such concepts as gravity and motion with his theory of four elements. He believed that all matter was made up of a combination of water, earth, air and fire, and a fifth element referred to as *aether*, which occupied the superlunary realm.

Aristotle believed that these four elements have a natural place and that the fundamental basis of motion in the sublunary realm is the tendency for these elements to return to that natural place. For example, if a stone is released, it moves downward toward the center of the cosmos, its natural place; flames rise upward toward the circumference, their natural place. Aristotle's ideas became

popular for many centuries in Europe, leading to scholastic and scientific developments in the Middle Ages and until the time of Galileo Galilei and Isaac Newton.

Aristotle's importance to the development of Western physics is due especially to its focus on explanatory power and coherence informed by observation. This drive toward coherence and explanation found expression in Aristotle's commitment to four modes of causation: material (that out of which a thing is made), formal (the archetype or plan), efficient (what it is that does the doing), and final (the end for which something is done). These four causes and their interpretation puzzled but also animated natural philosophers throughout the development of modern physics. (Typically today one says that physics concerns itself only with efficient causes.)

Heavenly Bodies and the Earth as Round

As with physicists since their time, the ancient Greek natural science philosophers were also interested in heavenly bodies. Anaxagoras was the first to learn that the Moon is illuminated by another light source; the Moon itself does not emit light. He also asserted that the Moon is much closer to the earth than is the Sun. He classified the Sun as a star because he believed that stars were large, luminous, and burning rocks. Anaxagoras believed that stars are further from Earth than the Sun because their brightness is not as strong.

In early Classical Greece, the idea that the earth is round was generally understood by everyone, and around 240 B.C.E., Eratosthenes accurately estimated its circumference. In opposition to Aristotle's geocentric views, Aristarchus of Samos (a Greek citizen who lived between 310 and 230 B.C.E.) presented an argument for a heliocentric model of the solar system, which placed the Sun at the center instead of the Earth.

A follower of this heliocentric theory, Seleucus of Seleucia asserted that the Earth rotated around its own axis, which also revolved around the Sun. Though his arguments have been lost in time, Plutarch stated that he was the first person to prove the heliocentric system through reasoning.

Archimedes

Archimedes of Syracuse, a Greek mathematician in the 3rd century B.C.E., is considered by many to be the greatest mathematician of antiquity and one of the greatest mathematicians of all time. He laid the foundation of statics and hydrostatics and calculated the mathematics of the lever. He also developed an elaborate system of pulleys that allowed one to move large objects with very minimum effort. The Archimedes screw is an important part of today's hydroengineering systems, and in his time, they helped the Roman army win the First Punic War.

Archimedes did not believe the arguments of Aristotle and metaphysics, believing rather that it was impossible to separate nature and mathematics, and he proved his assertions by converting mathematic theories into practical inventions.

In his work *On Floating Bodies*, which was created around 250 B.C.E., Archimedes discovered the law of buoyancy, which is also known as *Archimedes' Principle*. In mathematics, he utilized the method of exhaustion to calculate the area underneath the arc of a parabola with the summation of an infinite series. By doing this, Archimedes was able to give a rather accurate approximation of π .

Additional work by Archimedes includes the principles of equilibrium states and centers of gravity, formulae for the volumes of surfaces of revolution, and a system for expressing very large numbers. These theories would later influence Newton, Galileo, and Islamic scholars.

Hipparchus

Hipparchus, who lived from 190 to 120 B.C.E., had a life focused on mathematics and astronomy, and he utilized advanced geometrical techniques to map the motion of the planets and the stars, with the side effect of learning that solar eclipses would occur. Hipparchus also added calculations to determine the distance of the Moon and the Sun from Earth based upon improvements he made to observation instruments of his time.

Ptolemy

Ptolemy, who lived from 90 to 168 C.E., was another early physicist and was one of the leading

scientists of his time during the Roman Empire. He crafted several scientific treatises, at least three of which continued to be important throughout later European and Islamic science. The first of his astronomical treatises was known as the *Almagest* (or “The Great Treatise” in Greek—known originally as the *Mathematical Treatise*.) The second treatise was in the field of geography, which discusses the geographic knowledge of the Greco-Roman world.

Much of the information from Ptolemy and the rest of the ancient world was subsequently lost. Even the majority of the works from the greatest thinkers was lost with time. For example, though Hipparchus wrote at least 14 books, little of his direct work has survived. Only 30 of the 150 reputed Aristotelian works still exist today, and some of them are not complete.

India and China

Outside of Greece and Rome, specifically in the territories occupied by present-day India and China, there were also huge gains being made in physics.

Kanda, a philosopher from India, was the first to systematically develop a theory of atomism around 200 B.C.E. (although it is important to note that some historians have attributed him to the earlier era of 6th-century B.C.E.). This atomic theory was further elaborated by Buddhist atomists Dignaṅga and Dharmakīrti during the 1st millennium C.E. Another Indian philosopher by the name of Pakudha Kaccaya, a 6th-century B.C.E. Indian philosopher, also produced ideas about how the atom constructed the material world. All of the aforementioned philosophers believed that elements, except for aether, were physically palpable and were composed of very small particles of matter. The smallest of these particles (which could not be subdivided any further) would be called *Parmanu*. Atoms were believed to be eternal and indestructible.

Buddhists believe that atoms were not visible by the naked eye and that they could come and vanish in an instant, while the Vaisheshika philosophers believed that an atom was only a point in space. Indian theories were bathed in philosophy and were greatly abstract—they were not as grounded in research as the Greek and Roman philosophes.

They were based on logic—not on personal experience or experimentation.

In terms of Indian astronomy, the *Aryabhatiya* of Aryabhata proposed Earth’s rotation in 499 C.E., while Nilakantha Somayaji of the Kerala schools proposed a semi-heliocentric model that looked like the Tychonic system in the 1500s—long after the Greeks and Romans. Shen Kuo (1031–1095) wrote about Ancient China and magnetism in the *Book of the Devil Valley Master*.

The Role of Mathematical Theorizing in the West

The Pythagorean theorem is a relation in Euclidean geometry among the three sides of a right triangle. The theorem states that the side opposite the right angle (called the *hypotenuse*) is equal to the sum of the other two sides. The equation is written as

$$c^2 = a^2 + b^2$$

This is called the *Pythagorean equation*. The letter *c* represents the hypotenuse, and *a* and *b* represent the lengths of the other two sides.

The theorem is named after the Greek mathematician Pythagoras, who lived from 570 B.C.E. to 495 B.C.E. He is credited through tradition with the proof of the theorem, though many have said that knowledge of it predates him. There is evidence that Babylonian mathematicians understood the idea, though there is little surviving evidence that they used it as a mathematical formula. Chinese, Indian, and Mesopotamian mathematicians have all been credited with independently discovering the theorem. Some of these scholars even provided proofs of special cases.

Pythagoras is today one of the famous (and controversial) ancient Greek philosophers. He lived his early life on the island of Samos, off the coast of modern Turkey. At the age of 40, he emigrated to what is now southern Italy, where the majority of his philosophical activity occurred. However, Pythagoras wrote nothing, and there were no writings of his ideas recorded by others. By the first century of the Common Era (C.E.), Pythagoras was represented as a semi-divine figure in a largely unhistorical fashion. He was said, at the time, to have originated all true Greek

philosophy. A number of treatises were forged in his name to support this view.

We think of Pythagoras as a master scientist and mathematician. The early literature shows that he was famous in his own day—and not for mathematics. Instead, he was known as an *expert on the fate of the soul after death*. He believed that the soul was immortal and went through a series of reincarnations. He was also known as an *expert on religious ritual*, the founder of a life that emphasize religious ritual, intense self-discipline, and dietary restrictions. In addition, he was reputed to be a wonder worker who had a thigh of gold and could be in two places at once.

Today, controversy occurs around the idea that Pythagoras engaged in the rational cosmology that is considered indicative of the pre-Socratic scientists/philosophers. It was a matter of opinion whether he was any sort of mathematician. If early evidence is to be believed, he did present a cosmos that was structured according to moral principles and significant numerical relationships, which were similar to the ideas found in Platonic myths.

In this sort of cosmos, the planets were known as *instruments of divine vengeance* (such as “the hounds of Persephone”), and the Sun and Moon are known as *where people go when they live a good life* (the “isles of the blessed”). The thunder, for its part, was representative of a threat—used to frighten the souls being punished in Tartarus, the infernal region of the underworld.

It was also believed that the heavenly bodies appeared to have moved in accordance with the mathematical ratios that govern the musical instruments and produce music from heaven known as *the harmony of the spheres*. It is not believed that Pythagoras thought in terms of spheres. There is, however, evidence that he valued relationships between numbers (as in, e.g., the Pythagorean theorem).

Pythagoras’ idea of the cosmos was developed to be more mathematical/scientific by his successors, Archytas and Philolaus. Because of his optimistic argument about the fate of souls, Pythagoras succeeded in advocating a life that was attractive to his devoted followers.

That is not to say that Pythagoras did not believe in mathematical ways of thinking. He, together with his students, attempted to explain

the world utilizing numerical relations. Plato was mentor to Aristotle, and he believed that no wisdom could exist outside of mathematical reason.

Other followers of Pythagoras developed the theory that Earth was spherical and not located in the center of the universe. This theory stated that the Moon, the Sun, and the planets, including Earth, revolved around a central fire. This fire could not be spotted by anyone on Earth, however, because of an “anti-Earth” that blocked its view.

Geometry

The mathematical science of geometry, as used to understand nature, was first employed in both astronomy and physics in the 4th century C.E., maturing by the Hellenistic and Greco-Roman period. The Pythagorean mathematical point of view of nature was upheld during this time, but as the mathematical view was coupled with the world of the senses (as in the Platonic dialogues), little sound physics could be extracted. Though thinkers such as Plato were the leading theorists of their time, many of their ideas were later disproven. The physics of Aristotle continued to remain influential for centuries, although eventually it was shown that many of his ideas were incorrect.

One area in which early Greek mathematicians excelled was geometry, or the study of classical problems in mathematics. Three major problems arose, concerning the following mathematical issues: (a) angle trisection, (b) doubling of the cube, (c) constructing a cube which has a volume double that of a given one, (d) squaring a circle, (e) trisecting an arbitrary angle, (f) constructing a regular heptagon (which is a polygon with seven sides), and (g) constructing a square that has an area equal to that of a given circle.

These problems stood out for the ancient Greeks because they did not have solutions or their solutions were extraordinarily difficult to determine. The Greeks could not find simple solutions for these problems, and they were further impeded by the fact that all known solutions violated an important condition for these kind of issues—one of which the Greek mathematicians themselves had imposed. This condition stated that all valid solutions to construction problems are composed of a finite number of steps, which

consist of only two types: drawing a circle with a given center and radius and drawing a straight line through two points with a ruler (or a straight-edge). The fact that no solution exists subject to the self-imposed constraints was proved only rather recently—in the 19th century.

Doubling the Cube

Doubling the cube is the most famous of this collection of problems, and it is often referred to as the *Delian problem*, owing to the legend that the Delians consulted Plato on the matter. Another story states that the Athenians in 430 B.C.E. consulted the oracle at Delos to try and stop the plague their country was experiencing. These Athenians were advised by Apollo to double his altar that was created in the form of a cube. The scientists tried and failed several times to satisfy the god, and they eventually had to turn to Plato for advice to avoid Apollo's wrath. According to a letter from the mathematician Eratosthenes to King Ptolemy of Egypt, Euripides mentioned the Delian problem in one of his tragedies, which is now lost.

Constructing a Regular Heptagon

Regular squares, triangles, hexagons, and pentagons are all easily constructible. Octagons, on the other hand, are constructible on the heels of squares with a single angle bisection. All polygons obtained from the above four by doubling the number of sides are also constructible. However, this does not hold true for a heptagon, a seven-sided polygon.

By the year 1796, at the age of 19, Carl Friedrich Gauss (1777–1855) showed that a regular heptadecagon (a 17-sided polygon) is constructible. He showed that this heptadecagon was only a particular case of a family of constructive regular polygons. This is to say that an n -gon is constructible whenever n is a special kind of product of terms: The first is a power of 2, the rest (and there may be none) are distinct Fermat primes (that is numbers that are prime and of the form: 1 added to 2 itself raised to the power of a power of 2). Later, the French mathematician Pierre Wantzel proved that no other type of regular polygon is constructible.

Angle Trisection

Angle trisection is another of the Greeks' famous problems, though today we believe that

there is little difficult in trisecting an angle. In the classical framework, there are two types of lines drawn: circles and straight lines. The same simple approach applies to bisecting both circular arcs and straight-line segments.

The issue of trisecting a line segment is as simple as finding its n -th part for an arbitrary n . However, the issues of trisecting an angle (such as an arbitrary angle) is not solvable in a finite number of steps. Two points must be made, the first of which is that some angles are trisectable with a compass and ruler. One example of this is the three quarters of the straight angle: 67.5° . The angle is obtained by two consecutive angle bisections and makes the angle itself constructible. The third angle is obtained along the way. Angles of 45° (bisect a right angle) and angles of 30° (draw a right triangle with a side 1 and hypotenuse 2) are both constructible. The latter also leads to a classical trisection. This leads to some difficulty, as an arbitrary angle cannot be trisected with a ruler and a compass (as the Greeks asserted that it must be).

Another example is as follows: if one considers the geometric series $1/4 + 1/16 + 1/64 + \dots$ that adds up to $1/3$, we know that that adds up to $1/3$. Once an angle is bisected, we may also find its 2 - n -th part, for any n , and one may construct $1/4$, $1/16$, $\dots$ of any angle and, in principle, find its third after an infinite number of steps. This solution is universal, but it requires an infinite number of steps, which is forbidden.

This issue was solved in 1837 by Wantzel, who proved that there was no solution for trisecting a 60° angle in the classical framework.

Squaring a Circle

All of the solutions for the impossibility proofs show that each depends on demonstrating that roots of third-degree polynomial equations are not constructible. Squaring a circle, however, sits in a completely different category. The issue was solved in the 19th century, when the number π was shown to be transcendental. It is not algebraic, and it is not constructible. However, the area of a circle with a radius of 1 is exactly π . If a square with such area were constructible, the number itself would be, as well.

All of these mathematical puzzles and the limits to thought they reveal are interwoven with the

history of attempts to mathematize nature and thereby to provide clear and coherent explanations in terms of underlying processes of the phenomena we see around us.

Katie Moss

See also Physics, 19th Century; Physics, 20th Century; Physics, Contemporary; Physics, Enlightenment; Physics, Middle Ages; Physics, Renaissance

Further Readings

- Ashtekar, A. (2006). Space and time: From antiquity to Einstein and beyond. *Resonance*, 11, 4–19. <https://doi.org/10.1007/BF02834329>
- Bernal, J. D. (1997). *A history of classical physics: From antiquity to the quantum*. Barnes & Noble.
- Clagett, M. (1963). *Greek science in antiquity*. Macmillan.
- Jones, A., & Taub, L. (Eds.) (2019). *The Cambridge history of science: Volume 1, Ancient science*. Cambridge University Press.
- Weaver, J. H., Motz, L., & McAdoo, D. (1987). *The world of physics. A small library of the literature of physics from antiquity to the present*. Simon & Schuster.

PHYSICS, CONTEMPORARY

Theories are conceptual frameworks that play a leading role in articulating the achievements, state, and condition of contemporary physics. An understanding what the leading theories are, how they formed, how to interpret them, and their societal contexts is essential to the study of the philosophy and history of physics. This entry traces the path of theory in physics from its early development in the 16th and 17th centuries through the classical era of Newtonian mechanics and subsequent 20th-century development of relativity and quantum theory up to present-day interpretations and emerging theories such as string theory, supersymmetry, and others.

The term *theory*, however—even in contemporary mathematical physics—is ambiguous, and can have a gamut of meanings, including speculation, hypothesis, explanation, interpretation, and solidly or fully established package. As this entry

shows, the word can apply to a range of structures, for instance to a comprehensive account of a single comprehensive area of physics, as *quantum theory*, or to a subset thereof, as quark theory. The word has still other kinds of uses; the expression “effective field theories” refers to structures that seek to model effects without modeling causes, such as in BCS theories of superconductivity (named for its formulators John Bardeen, Leon Cooper, and Robert Schrieffer). But far from being a weakness, these ambiguities are a strength, part and parcel of why theories have been and continue to be so important in physics. Some proposed theories, for instance, initially have had no experimental tests, or have even been experimentally disproven, and only later have become indispensable, without it ever being incorrect to refer to them as theories.

Several factors make a theory robust despite these many ambiguities. First, theories *frame* existing knowledge, organizing what scientists have learned, giving them a grip on what is possible and impossible at the present moment. Electromagnetic theories, for instance, systematize present-day knowledge about electric and magnetic phenomena. Second, theories *scale* such knowledge, showing how to relate what happens at one level to what happens at another, the way that electromagnetic theory allows us to relate the behavior of charged bodies to that of subatomic and astronomical phenomena. Third, theories *orient* scientists by indicating how to utilize and explore phenomena such as electromagnetism, indicating what questions need to be investigated and suggesting ways to go about investigating them. Because theories frame, scale, and orient, they possess a coherent core that is more important than more external properties such as empirical confirmation or testability. It is therefore wrong to refuse to deem something a theory simply because, say, it collides with experimental results or may be currently untestable.

One notable example of a theory that contradicted existing knowledge is Yang–Mills theory, a mathematical scheme published in 1954 by Chen Ning Yang and Robert Mills, as a possible tool for describing the strong interaction in elementary particle physics. It was in direct conflict with at least one critical element of what was known

about particle physics at the time. In existing particle physics theory force-carrying particles had to be massive, while in the proposed Yang–Mills theory, they were massless. The authors published their theory anyway. Two decades later, the modification of other elements in particle physics theory made it not only possible but necessary to incorporate Yang–Mills theory into particle physics, and today, Yang–Mills is one of its core structures.

Another example is string theory, in which elementary particles are not viewed not as points but as one-dimensional strings. It makes no predictions that are experimentally testable now or that will be in the immediate future, and for this reason, some physicists have denounced it—albeit caustically—as having a similar epistemological status as astrology, intelligent design, and other forms of pseudoscience. Nevertheless, many theoretical physicists not only work in string theory but also consider it one of the most promising approaches in physics, as potentially reconciling relativity and quantum mechanics.

Historical Origins

The roots of contemporary theories of physics began to develop in the late 16th and early 17th centuries and sprang from the confluence of at least three key historical and cultural developments in particular.

One of these was the way that nature came to be seen as occupying its own realm, separate from that of the spiritual. According to early Catholic doctrine, nature was Creation, and to be understood as the product of divine activity, and not therefore to be completely distinguished from scriptural writings and the divine mind. “The heavens declare the glory of God,” reports Psalm 19 in the *Bible*, “and the firmament sheweth his handywork.” One traditional way that early theologians expressed this was to say that God wrote two books, the Book of Scripture and the Book of Nature, which had to be read together as they shed light on each other. By the 16th century, however, several historical developments made it seem as if these two books had no connection and had to be read separately. These developments included numerous voyages of adventure in which explorers turned up things not discussed in the Bible, and

the Protestant theological insistence that the Bible can and must be read by individuals relying on its words alone.

How, then, *did* one read the Book of Nature? The second key development was that the syntax of this book was mathematical. This development is identified most closely with Galileo Galilei, who argued—rhetorically, at least in part—that the Book of Nature was written in mathematical figures, and that thus only those who understood mathematics were able to read it. Still, a third intersecting development was the growing influence of theory-making, or the belief that true knowledge required stepping back and viewing what is to be known from a distance, as if one were looking at those things on the other side of a glass plate.

Classical Theories

These three developments form the background for the development of classical mechanics in the 17th century. The foundation of classical mechanics is provided by Isaac Newton’s 1687 book *Philosophiae Naturalis Principia Mathematica*, known simply as the *Principia*, whose definitions, axioms, and proofs outlined a theory of the laws of motion and of universal gravitation. Classical mechanical theory is based on several assumptions. One is that of homogeneity: space, time, and laws apply across all scales, from the atomic to the cosmic. Another is that of *simple location*, or that things to which it applies have distinct identities and are always located at a particular place at a particular time. Yet other assumptions are that the movements of these things are deterministic and predictable and due only to mechanical pushes and pulls without requiring to be explained with reference to human thought and actions.

In Newtonian theory, the world is conceived as akin to a fishbowl that contains only bodies, the only thing these bodies do is move, and the only things that make them move are forces—and that we peer into this world from the outside. The world to which classical mechanical theory applies is a mechanical, mathematical, and thoroughly comprehensible, in which all objects—grains of sand, chandeliers, swings, cannonballs, dancers, planets, and galaxies—are treated as objects moving under the influence of forces, and whose

motions therefore can be calculated and predicted. Newtonian theory therefore promised that the world was fully intelligible, and it considerably broadened the ability of humans to dominate and control it. Its success at this has tempted some philosophers into instrumentalism, or the view that the purpose of theories is to dominate and control nature.

Historians of science run the gamut from pure externalists, who approach all developments in science—including theories—as social products, to internalists, for whom theories are to be viewed as arising in the lab from scientists motivated by data. An extreme externalist was the Marxist Boris Hessen, who in his 1931 article “The Social and Economic Roots of Newton’s *Principia*” treated classical mechanics as principally the successful product of bourgeois members of industrialized society who were seeking to dominate nature and control the working classes through improvements in the modes of production and in weaponry. *Internalist* scholars, however, have attacked the *Hessen thesis*, arguing much more plausibly that Newton’s work arose principally from his understanding of predecessors and his desire for coherence and universality.

The laws, principles, and axioms of classical theory had a huge social impact in areas as diverse as education, religion, literature, political theory, and philosophy. It inspired early socialists as well as Karl Marx to seek deterministic and universal laws, and it inspired the framers of the U.S. Constitution as well. By seeming to outline criteria of objectivity, classical Newtonian theory also strongly influenced numerous philosophers, including David Hume and Immanuel Kant. So strong was its social impact that Newtonian theory has been described as the foundation of modernity. Newtonian theory persisted for over 300 years with its basic outlook virtually intact. New phenomena, concepts, and theories appeared—such as interference, wave phenomena, and the Maxwellian theory of electromagnetism—but nearly all of their elements were incorporated into classical mechanical theory.

Then came two bombshells at the beginning of the 20th century, relativity and quantum mechanics, which exposed elements of classical theory as mere assumptions, and which required new kinds

of theories and whose scientific and social effects were extensive. For all their considerable differences, both relativity theory and quantum theory coincidentally developed in two phases, with the first phase of each having minor, and the second major, scientific and social impact.

Relativity

Unlike quantum mechanics, the development of relativity theory in each phase came about almost completely through the efforts of a single person, Albert Einstein. Its first phase, known as *special relativity*, was the product, in brief, of Einstein’s attempt to resolve one of the few inconsistencies that arose between two previous theoretical systems, Newton’s and that of the 19th-century Scottish mathematician James Clerk Maxwell. At the core of classical Newtonian mechanics is the relativity of motion, meaning that the laws of physics stay the same in any inertial reference frame, or a reference frame moving at constant speed. The same laws of physics, that is, describe billiard balls, falling objects, and the movements of dancers on the ground—and on a smoothly moving train or on any platform moving constantly and linearly at any speed. In an inertial reference frame, that is, no experiment could be carried out to determine whether or not it was in motion. The equations used to connect motions in two different inertial reference frames are called *Galilean transformations*. But according to Maxwell’s theory of electromagnetism, light moves at constant speed regardless of reference frame. Light was assumed to have this property because it was assumed, like sound, to move in a medium that fixed its motion, a medium called *ether*. If light did, however, it should travel at different speeds when moving in different directions. In experiments conducted in 1881 and 1887, the physicists Albert Michelson and Edward Morley failed to detect such a difference; light indeed seemed to travel at the same speed regardless of inertial frame. The special theory of relativity, which Einstein proposed in a 1905 paper, was the product of an attempt to develop sets of equations, known as *Lorentz transformations* (named after the Dutch physicist Hendrik Lorentz), to reconcile the theories of Newton and Maxwell, without invoking effects of the ether, by giving the compensations of space and time needed

to preserve the constancy of the speed of light in different inertial frames of reference. In another 1905 paper, Einstein drew out a dramatic consequence—that the mass of a body is related to the energy it contains, by the famous formula $E = mc^2$, plus a less famous corrective factor (omitted in popular presentations) related to the body's motion.

Whereas the theory of special relativity arose from preserving invariance in inertial reference frames, general relativity arose from preserving invariance in accelerated reference frames. Einstein proposed, for instance, that no experiment could determine whether one were inside an elevator car accelerated by forces (or floating in empty space) or at rest in a gravitational field (or freely falling in one). He now sought another set of equations to describe what happens in an accelerated frame of reference. The implications of this set of equations, called *general relativity*, has profound consequences for space and time, and makes the general theory of relativity radically different from classical theory and even special relativity. Space and time are linked, general relativity implies, and the resulting space-time is curved by the mass in it. Space-time tells mass how to move, in a famous formulation, and mass tells space-time how to curve. This theory also requires a fourth-dimensional surface to represent the gravitational field.

Revolutionary new scientific theories rarely spring solely from laboratory work, and several scholars have proposed that certain social influences may have played a contributing role to the development of relativity theory—that the pressing problem in Europe at the beginning of the 20th century of coordinating time in order to make trains arrive, depart, and run on schedule helped inspire special relativity, and that the growing awareness of curved spaces in non-Euclidean geometry and of fourth-dimensional space were at work in the development of both relativity and modern art.

General relativity had a strong impact on philosophy, principally via its implication that the shape of space is not inherently Euclidean, as nearly all philosophers, including Kant, had assumed; Kant, in fact, had claimed that the *form* of human sensibility was transcendental—non-empirical—and necessarily Euclidean. But according to the theory of general relativity, the shape of space was not intuitive but to be empirically determined. This hardened a rift between two groups

of philosophers: logical empiricists, who saw the theory of general relativity as refuting Kant and as indicating that sensibility and the understanding were split, and transcendental philosophers, who saw general relativity as mandating a revision of the necessary link that Kant saw between sensibility and the understanding.

Quantum Mechanics

In 1900, Max Planck proposed what is now known as the Planck energy formula, $\epsilon = nh\nu$. One of the most fundamental equations expressing the structure of the world, it says that the energy (in Planck's initial case, of resonators of a certain frequency) could not have any energy value, as classical physics implied, but only values separated by an integer multiple n of $h\nu$, ν being the frequency, and h being what has come to be known as Planck's constant, a measure of the smallest quantum of energy. Planck knew that what he was doing was inconsistent with classical theory, and introduced it as a stop-gap measure, on the assumption that his proposal could be eventually reconciled with classical theory. He did not imagine that the idea, introduced to explain puzzling experimental results, would revolutionize physics theory in a way that would eventually replace classical theory.

During the quarter century that followed, however, new developments inserted the quantum ever more fundamentally into physics theories: Niels Bohr introduced it into the structure of the atom in 1913–1914, Einstein used it in papers of 1916–1917 that put probability into the foundations of atomic theory, and it was an element of Wolfgang Pauli's exclusion principle, proposed at the beginning of 1925. During this quarter-century, however, quantum theory remained essentially a collection of amendments to classical theory; quantum theory, so to speak, was like having only a set of instructions for getting from one place to another when what one really needed was a map of the entire terrain. Meanwhile, the public viewed quantum theory as akin to relativity in that, however important and fundamental, it was unintelligible to the layperson. Quantum theory therefore had very little cultural impact, though it did introduce a few terms into public vocabulary. One was *quantum leap*, which in an age of typewriters and

telegraphs was warmly embraced to refer to tiny discontinuous transitions but has since come to refer to gigantic transitions.

Then, in one of the most startling and far-reaching episodes in the history of science, in the short space of the 2 years 1925–1927, not one but two such maps appeared. One was created by Werner Heisenberg, who had assumed that what was blocking physicists from developing a full theory of the quantum world was their assumption that its events were picturable (*anschaulich*), when in fact quantum phenomena were unpicturable, *unanschaulich*. Quantum phenomena, he proposed, could be described only by a special form of mechanics, called *matrix mechanics*, in which certain properties were not commutable. The second map, created by Erwin Schrödinger, was based on the much more picturable wave theory, but whose waves were about the *probabilities* of events. Although sharply different, the two maps were soon shown to be essentially identical, in what became known as *quantum mechanics*.

The newly mapped territory was much different from Newton's. It was marked by inhomogeneity, or differences across scales, applied to the quantum and to the classical worlds. It also included violations of *simple location*, for it included *entangled* phenomena, whereby where states are linked despite being physically far apart—unheard of and impossible in the classical world. Some of its events were also unpredictable and uncertain, for certain properties could not be pinned down or even said to *exist*. Finally, however humans measure quantum phenomena, these phenomena cannot be viewed as if through a pane of glass but the act of measuring changes key properties of the phenomena being measured.

Quantum mechanics had a huge impact on human thought, one major example being the multiple philosophical and metaphorical valences attributed to Heisenberg's *uncertainty principle*. The uncertainty referred to is, in philosophical vocabulary, not just epistemological, or having to do with what we *know* about nature, but *ontological*, or having to do with nature itself. It refers not to information that investigators don't happen to know, but to information that in certain circumstances investigators *cannot* know. Quantum mechanics has had an impact on other fields. Within two years of the introduction of the

uncertainty principle, for instance, the philosopher John Dewey cited it as a confirmation of his spectator theory of knowledge. Quantum mechanics has contributed widely used images and terms in popular culture, such as complementarity, Schrödinger's cat, and alternate worlds.

Externalist historians and sociologists have occasionally attempted to account for quantum mechanics as a social product, with the most notorious of these the *Forman thesis*. In his 1971 article "Weimar Culture, Causality and Quantum Theory, 1918–1927: Adaptation by German Physicists and Mathematicians to a Hostile Intellectual Environment"—which is one of the most influential and controversial article in the historical studies of science—the historian of science Paul Forman argued that the founders of quantum mechanics were motivated to compose the acausal and probabilistic theory of quantum mechanics because of the anxiety of the post-World War I intellectual climate, which rebelled against rationalism and causality. As Forman saw it, quantum mechanics belonged to the same intellectual atmosphere that produced Oswald Spengler's influential 1918 book *The Decline of the West*. Internalists, however, have pointed out problems in the article, including its sources, which are mainly popular speeches rather than laboratory notebooks, as well as the fact that Einstein's introduction of probability into the foundations of the theory preceded Forman's start date of 1918.

Interpreting Quantum Mechanics

Although quantum mechanics is a complete, exhaustive, and thoroughly tested theory, its interpretation—how its terms and structures connect with the rest of human knowledge and experience—is one of humanity's great intellectual challenges. Its various interpretations have such radically different implications for the structure and character of the world that they are also themselves often called *theories*.

Copenhagen Interpretation

The first and most widely known interpretation of quantum mechanics, the *Copenhagen Interpretation*, was crafted by Niels Bohr and Werner Heisenberg in the late 1920s and early

1930s. It was named after Bohr's birthplace, as well as the site of his Institute of Theoretical Physics, where much of the theory was worked out. The Copenhagen Interpretation presents the world as the product of two ingredients. The first is an information function, the ψ -function described by Schrödinger's equation, which is a classical wave that expands outward and overlaps, or *superposes*, numerous possibilities. The second ingredient is a measurement or some other interaction with the classical world that evaporates or *collapses* the function, making one hitherto only probable state become *real* and eliminating the others. According to this interpretation, all we can ultimately know of the world prior to acts of observation or measurement is a set of probabilities, specified by the ψ -function, for which states might be realized. The Copenhagen Interpretation in effect divides the world into two entirely separate domains: the quantum world governed by the unobservable and intangible ψ -field and the deterministic classical world. As Bohr saw it, the Copenhagen interpretation was an effectively modified Kantianism: We must use classical language to describe as best as we can something to which classical language does not apply. The Copenhagen Interpretation challenged conventional notions so directly and deeply that it sparked a kind of philosophical opposition, even among scientists themselves.

Hidden Variables

One way to eliminate the division between the classical and quantum worlds is to deny the fundamental character of the first ingredient mentioned earlier, the wave function. This path was promoted most notably and persistently by Einstein, who insisted that quantum states don't represent real states of affairs but only states of our knowledge. In this interpretation, there is no true information wave function, no superposition, and no intrinsic probability of states appearing; quantum mechanics is simply incomplete. More factors exist than we have discovered, and more thought and inquiry will surely discover these and restore causality, determinism, picturability, and the other key features of the Newtonian world. Versions of the hidden variable theory

appeared at the very beginning of quantum mechanics, and have persisted to this day. But as all the predictions of quantum mechanics, even the strangest, have passed experimental tests, this theory survives on a promissory note.

Alternate Worlds

Another approach that rejects the Copenhagen interpretation denies the second ingredient mentioned above, the collapse of the wave function. Measurement and other contact with the classical world does not, and cannot, eliminate the other possibilities, which continue to exist, branching off from each other. Introduced in the 1950s, this interpretation also still has advocates. In restoring consistency and coherence to the theory, however, it does so by muddling the concept of reality.

QBism

Still another interpretation is known as *Quantum Bayesianism*, now QBism. Like Einstein, it says quantum states don't represent real states of affairs but only states of our knowledge; unlike Einstein, it embraces this, saying that this indicates not the incompleteness of quantum mechanics but that the probabilities derived from a quantum state are subjective, *Bayesian* probabilities, degrees of belief or uncertainty in a subject. In the context of quantum mechanics, QBism sees probabilities such as those of the Born rule (which gives the probability that a measurement of a quantum system will yield a given result) as normative rather than prescriptive, a guide for making judgments about quantum events. This interpretation eliminates many paradoxes. The *collapse* of the wave function simply updates the subject's beliefs. Entanglement disappears because the correlations are not simultaneously experienced. Because QBism starts from the experience of the subject, the approach bears certain affiliations with phenomenology.

But these are only a few interpretations of quantum mechanics; other, more technical interpretations include operationalism, decoherence theory, and the consistent-histories interpretation. Another set of theories ambitiously attempts to incorporate relativity theory and quantum mechanics; besides

string theory, these include M-theory, supersymmetry, supergravity, loop quantum gravity, relational quantum mechanics, and others.

Theory and Practice

Theories exist throughout branches of physics, often as a blend of quantum and classical theories. Sometimes, they exert more of an impact than others. The theory of defects, for instance, played a major disciplinary role in bringing together the physics of specific materials—metals, ceramics, glass, and so forth—into a single field, materials science.

Whereas theories play a leading role in articulating the achievements of contemporary physics, they may play less of a significant role in its practice. For in the laboratory environment, physics does not unfold abstractly independently of artifacts, tools, and practices, and it is therefore not free of instrumental mediation. Such forms of mediation are not just artifactual but also include analogies, images, and calculational tools including what Ursula Klein calls *paper tools*, which can shape the projects that theorists undertake and even imagine doing. The influence of these mediating influences may call into question the primacy of the role theories play in the practice of theorists.

Robert P. Crease

See also Physics, 19th Century; Physics, 20th Century; Physics, Quantum Theory; Statistical Inference, Bayesian

Further Readings

- Crease, R. P., & Goldhaber, A. S. (2014). *The quantum moment*. W. W. Norton.
- Dobbs, B. J. T., & Jacob, M. C. (1985). *Newton and the culture of Newtonians*. Humanities Press.
- Einstein, A., & Infeld, L. (1938). *The evolution of physics: The growth of ideas from early concepts to relativity and quanta*. Cambridge University Press.
- Feingold, M. (2004). *The Newtonian moment: Isaac Newton and the making of modern culture*. Oxford University Press.
- Forman, P. (1971). Weimar culture, causality, and quantum theory, 1918–1927: Adaptation by German physicists and mathematicians to a hostile intellectual environment. *Historical Studies in the Physical Sciences*, 3, 1–115.
- Galison, P. (2004). *Einstein's clocks, Poincaré's maps: Empires of time*. W.W. Norton.
- Graham, L. R. (1985, November). The socio-political roots of Boris Hessen: Soviet Marxism and the history of science. *Social Studies of Science*, 15(4), 705–722.
- Gribbin, J. (2015). *Einstein's masterwork: 1915 and the general theory of relativity*. Icon Books.
- Harrison, P. (2006). The bible and the emergence of modern science. *Science & Christian Belief*, 18(2), 115.
- Jammer, M. (1974). *The philosophy of quantum mechanics: The interpretations of quantum mechanics in historical perspective*. Wiley & Sons.
- Kaiser, D. (2005). *Drawing theories apart: The dispersion of Feynman diagrams in postwar physics*. University of Chicago Press.
- Kaiser, D. (2011). *How the hippies saved physics: Science, counterculture, and the quantum revival*. W. W. Norton.
- Kaiser, D. (2020). *Quantum legacies: Dispatches from an uncertain world*. University of Chicago Press.
- Pais, A. (1986). *Inward bound: Of matter and forces in the physical world*. Oxford University Press.
- Peterson, M. A. (2011). *Galileo's muse: Renaissance mathematics and the arts*. Harvard University Press.
- Rickles, D. (Ed.). (2008). *The Ashgate companion to contemporary philosophy of physics*. Ashgate.

PHYSICS, ENLIGHTENMENT

The Age of Enlightenment, which historians generally date from the mid-17th century through the 18th century, is a period in the history of Western thought and culture characterized by dramatic changes in philosophy, science, politics, and the social order. Increasing dissatisfaction with the medieval worldview and its domination by religious dogma and hereditary monarchies had already resulted in widespread political dissent, the Protestant Reformation, and decades of often violent conflict between adherents of the old order and militant reformers.

Among the great upheavals of the 18th century was the French Revolution, which followed in the wake of the successful American War of Independence against England, and in which the political power of the Catholic Church and the social orders of the French nobility and monarchy were called into question and then violently dismantled in the name of liberty and equality for all—ideals that both American and French reformers

claimed were founded upon the principles of human reason.

Ideas which were dismissed with the Enlightenment include many of the presuppositions that had long guided, and constrained, philosophical inquiry. Philosophy now gained the status of an independent force with the power to challenge old beliefs and construct new ones in the realms of both practice and theory, on the basis of its own principles. The revolution in science began when natural philosophers mounted an attack on the medieval Scholastic philosophical program and began to integrate concepts such as mechanics and astronomy, eventually leading to a universally valid characterization of motion and various other concepts.

The new branch of science was used to explain the natural world by accounting for a wide variety of phenomena with a relatively small number of mathematical formulae. The philosopher, mathematician, physicist, and music theorist Jean le Rond D'Alembert (1717–1783) was a leading figure of the French Enlightenment, and he wrote of it as the century of philosophy *par excellence* because of the incredible intellectual progress of the age, the enthusiasm of that progress, and the advances of the sciences. There was an idea that philosophy would improve overall human life.

During the 16th and 17th centuries, considerable scientific progress was made throughout Europe. The revolution in science began when natural philosophers created an attack on the Scholastic philosophical program and began to integrate concepts such as mechanics and astronomy, which led to a universally valid characterization of motion and various other concepts. Beginning in the 18th century, the mechanics first presented by Isaac Newton were developed by several mathematicians as they learned calculus, and its initial formulation was elaborated upon. Mathematical analyses were performed and applied to problems of motion. This was known as *mixed mathematics* (or rational mechanics) and was later called *classical mechanics*.

The following sections discuss important developments in the field of physics, among which are the predicted periodicity of Halley's comet by Edmond Halley in 1705; Benjamin Franklin's assertion in 1752 that lightning is electricity; the proposal of the electrical inverse-square law in 1767 by

Joseph Priestley; William Herschel's discovery of the planet Uranus in 1781; John Michell's assertion in 1783 that some objects are so massive that not even light could escape from them; the introduction of the inverse-square law of electrostatics in 1785 by Charles Coulomb; Henry Cavendish's measurement of the gravitation constant (which led to the determination of the mass of Earth) in 1798; and Count Rumford's idea that heat was a form of energy, which he proposed in 1798.

Physics, as Demonstrated Through Mechanics

During the Enlightenment, rational mechanics revolved primarily around the development of elaborate mathematical treatments of observed motions. These mechanics used Newtonian principles as a basis and emphasized improving the tractability of complex calculations, as well as developing a working means of analytical approximation. A textbook on these matters was published by the Hungarian professor of physics and philosophy Johann Baptiste Horvath (1732–1799) during this time, and by the end of the century, scholars had developed analytical treatments that were strong enough to verify the stability of the solar system based only on Newton's laws—divine intervention had no place in these theories. However, deterministic treatments of the three-body problem in gravitation and other simple issues remained intractable.

Significant Contributors

Among the Enlightenment's illustrious thinkers are Colin Maclaurin, who discovered uniformly rotating self-gravitating spheroid in 1742 and Leonhard Euler, who solved the ordinary differential equation for a forced harmonic oscillator, along with the phenomenon of resonance in 1739. However, British work, performed by mathematicians such as Maclaurin and Taylor, fell behind Continental developments, where academics were succeeding as led by mathematicians such as Euler, Pierre-Simon Laplace, Daniel Bernoulli, Joseph-Louis Lagrange, and Adrien-Marie Legendre.

In 1743, d'Alembert published his work *Traité de Dynamique*. Within this text, d'Alembert introduces the idea of generalized forces that can

accelerate systems and systems within constraints. Three years later, Pierre Louis Maupertuis applied minimum principles to mechanics. Shortly after, in 1759, Euler solved the partial differential equation of the vibration of a rectangular drum. He then found the Bessel function solutions for the differential equation for the vibration of a circular drum.

Other key thinkers include John Smeaton, who in 1776 published a paper on experiments that connected kinetic energy, work, momentum power and supporting the conservation of energy, and LAGRANGE, who presented his equations of motion in *Mécanique Analytique* in 1788.

In 1789, Antoine Lavoisier stated the law of conservation of mass. Newton's mechanics were discussed in both the *Celestial Mechanics* (1799–1825) of Laplace and Lagrange's 1788 work.

In 1733, the fundamental frequency of a stretched vibrating string (in terms of tension and mass per unit length) by solving a differential equation was derived by Brook Taylor. Following on this study, Swiss mathematician Daniel Bernoulli (1700–1782) studied and made discoveries concerning the behavior of gases, which anticipated the kinetic theory of gases proposed more than a century later. Bernoulli has often been called the *first mathematical physicist*. In 1734, 1 year after harnessing the power of the differential equation, Bernoulli solved the differential equation for the vibrations of an elastic bar clamped on one end. His examination of fluid dynamics and research of fluid flow were outlined in his work *Hydrodynamica*, published in 1738.

Thermodynamics

The field of thermodynamics was developed during the 18th century through the theories of weightless *imponderable fluids*. These included electricity, heat (*caloric*), and phlogiston (which was discarded as a concept when oxygen gas was identified late in the century). These fluids, as assumed real fluids, could be traced by their flow through chemical reactions or a mechanical apparatus.

Experiments in thermodynamics led to the creation of new types of experimental apparatus, such as the Leyden jar, invented in 1745 and used for storing static electricity. New types of measuring instruments were developed as well, including

the calorimeter. Improved versions of existing tools, such as the thermometer, were also developed, along with new concepts, such as Joseph Black's idea of latent heat and Benjamin Franklin's argument that electrical fluid flows between places of deficit and excess (later referred to as *positive* and *negative charges*). Franklin is most well known, however, for his discovery in 1752 that lightning is a form of electricity.

In the 18th century, as mentioned above, heat was viewed as a type of fluid, called *caloric*. This theory was later shown to be incorrect, but scientists of the time working under this assumption were able to make several important discoveries. These scientists included Henry Cavendish and Joseph Black. This theory of heat was opposed to the theory of heat that arose during Newton's time, which stated that heat is a product of the motions of the particles of a substance. This theory gained support in 1798 due to the experiments of Count Rumford (also known as *Benjamin Thompson*), who found a direct relationship between mechanical energy and heat.

Although scientists in the 18th century knew that finding absolute theories of electrostatic and magnetic force (much like Newton's principles of motion) would be an incredible achievement, none were easily found. However, as additional experimental practices became widespread and refined in the early years of the 19th century and analytical methods of rational mechanics were applied to experimental phenomena, especially with Joseph Fourier's analytical treatment of the flow of heat, a solution began to emerge. In 1767, the English polymath Joseph Priestley proposed an electrical inverse-square law, and, in 1798, Charles-Augustin de Coulomb introduced the inverse-square law of electrostatics. (Both Priestley and Coulomb are discussed later in this entry.)

By the end of the 18th century, the members of the French Academy of Science were clearly leaders in the field of mechanics; however, the traditions established by Galileo Galilei and his followers continued to exist. Along with the French Academy of Science, the Royal Society, founded in London in 1660, was a major center for the performance and reporting of experimental work in the mechanics, as well as in the fields of static electricity, optics, magnetism, physiology,

and chemistry. As these areas became more defined, many advancements were made in each field. Chemical experimenters, for example, focused on the classification and isolation of chemical reactions and substances instead of focusing on Newtonian chemical affiliations.

Astronomy: Halley's Comet

Halley's Comet, also referred to as *Comet Halley* or *IP/Halley*, is the best known of the short-period comets. It is visible from Earth every 75 or 76 years with only the naked eye. Halley is the only naked-eye comet that one might see twice in their lifetime. It last appeared to the inhabitants of Earth in 1986 and will next appear in mid-2061.

The comet's habitual returns to our inner solar system have been seen and recorded by astronomers since at least 240 B.C.E. Instances of the comet's appearance have been recorded by the ancient Babylonians, the Chinese, and medieval Europeans, but at the time, the instances were not connected as sightings of the same comet. Its periodicity was not determined until 1705, when English astronomer Edmond Halley made the discovery.

In 1986, when the comet was last visible from Earth, it also became the first comet to be observed in detail by spacecraft. This allowed it to provide the first observational data on the structure of a comet's nucleus and to support a number of longstanding hypotheses about how comets were constructed. Perhaps most notable was the American astronomer Fred Whipple's *dirty snowball* model, which correctly asserted that Halley was composed of a mixture of volatile ices, such as carbon dioxide, water and ammonia—as well as dust. The mission also evolved and improved these ideas. For example, we now know that only a small portion of Halley's Comet is icy—most of it is *dusty*.

Halley's Comet has likely been in its present orbit between 16,000 and 200,000 years, though it is not feasible to numerically calculate its orbit based on the few recorded observations we have. We do know that when this comet approaches the Sun, it expels gas from its surface, and it is knocked slightly off its path. This causes delays of about four days each time it happens. It is thought that comets of the Halley type were originally long-period comets whose orbits were disturbed by the

gravity of the giant planets—at which point they were sent into the inner Solar System. If Halley was indeed of this type, it is likely to have originated at the Oort Cloud, though others have proposed that it came from the area near Uranus and Pluto.

Halley's Comet was the first to be recognized as periodic. Until the Renaissance, it was accepted that Aristotle and his followers had been correct in asserting that comets were disturbances in Earth's atmosphere. This idea was disproven, however, by the Danish astronomer Tycho Brahe in 1577. He used parallax measurements to prove that comets must also occur beyond the Moon. At the time, most of the world did not believe that comets orbited the Sun—instead they believed that they followed straight paths through the solar system.

This is where Newton's friend Edmond Halley comes in. Halley, who was a publisher and editor, published his *Synopsis of the Astronomy of Comets* in 1705. In this work, he used his friend Newton's laws to calculate the gravitational effects of Saturn and Jupiter on cometary orbits. After making this calculation, Halley was able to determine that the orbital elements of a second comet that had appeared in 1682 was the same comet that had appeared in 1531 (observed by Petrus Apianus) and 1607 (observed by Johannes Kepler). He noticed that the same comet returned every 76 years, which we now know to be every 75–76 years. That year, Halley correctly predicted the comet would make its way back to Earth in 1758.

Johann Geor Palitzsch, a German farmer and amateur astronomer, saw Halley's Comet on December 25, 1758. The comet did not pass through its perihelion until March 13, 1759, because Saturn and Jupiter caused a retardation of 618 days. A team of three French mathematicians calculated this effect before its return (with a 1-month error). These mathematicians were Joseph Lalande, Alexis Clairaut, and Nicole-Reine Lepaute. Halley, unfortunately, died before he could view the comet's return for himself. The comet's return also represented the earliest successful tests of Newtonian physics. French astronomer Nicolas Louis de Lacaille named the comet in Halley's honor in 1759.

It is important to note that though Halley's Comet was named after Edmond Halley, others

may have noticed its appearance long before he did. The idea has been raised that first-century Jewish astronomers knew the comet was periodic, according to a passage in the Talmud.

Benjamin Franklin and the Lightning Rod

Benjamin Franklin (1706–1790) is known as one of the founders of the United States and as the inventor of a number of devices that are still in use today. Benjamin Franklin also invented the pointed lightning rod conductor, which was also called a *Franklin rod* or *lightning attractor*. This invention occurred in 1749 and was part of Franklin's pathbreaking explorations of electricity. Although not the first to write about the correlation between lightning and electricity, Franklin was the first to propose a workable system for testing his hypothesis. It was several more years before Franklin performed his reported kite experiment, which, through legend, took place because he was tired of waiting for Christ Church in Philadelphia to finish being built so he could put a rod on top of its spire. (It is important to note, however, that the actual occurrence of this event has been heavily debated).

Later, in the 19th century, the lightning rod was used as a decorative motif. They were embellished with ornamental glass balls and are now a popular collector's item. The balls have been used as weather vanes, but their main purpose is to provide evidence of a lightning strike by falling off or shattering. This helps building owners know if they should check their buildings and ground wires for damage after a storm. The marine lightning rod was also developed soon after Franklin's initial work.

Joseph Priestley

The Englishman Joseph Priestley (1733–1804) was an 18th-century dissenting clergyman, theologian, chemist, natural philosopher, Liberal political theorist, and educator who published over 150 works. Priestley is typically credited with the discovery of oxygen, as he isolated it in its gaseous state. However, Antoine Lavoisier and Carl Wilhelm Scheele also have claims to the discovery.

Throughout his life, Priestley's considerable scientific reputation was contingent on his invention of soda water, his discovery of several *airs* (gases), and his writings on electricity. He named oxygen *dephlogisticated air* at the time. He was determined to defend phlogiston theory and reject what would become the chemical revolution, but this eventually left him isolated within his own scientific community.

Priestley also saw a strong connection between science and his brand of theology. He made repeated attempts to connect Christian theism and Enlightenment rationalism. He attempted to combine determinism, theism, and materialism in his metaphysical texts. He thought that a good understanding of the natural world would promote human progress and bring about a Christian Millennium.

Priestley strongly believed in an open exchange of ideas, and he advocated toleration and equal rights for religious Dissenters. This also led to his cofounding of the Unitarian Church in England. His controversial publications and outspoken support of the French Revolution aroused suspicion from the government and the public. In 1791, he was forced to flee from Birmingham with his family. First the Priestleys went to London in 1791, then fled to the United States, following a mob's burning down his church and home. In the last 10 years of his life, he dwelled in Northumberland County, Pennsylvania.

In addition to his contributions to science, Priestley contributed to pedagogy, publishing works on English grammar, early and influential timelines, and books on history. In fact, his educational works were some of Priestley's most popular. However, his metaphysical works did outpace them. Leading philosophers such as Herbert Spencer, John Stuart Mill, and Jeremy Bentham later credited them as some of the primary sources for utilitarianism.

Charles Augustin de Coulomb

Coulomb's law, also called Coulomb's inverse-square law, describes the electrostatic interaction between electrically charged particles. First published in 1785 by the French physicist Charles-Augustin de Coulomb, the law was essential to the development of the theory of electromagnetism. It is also analogous to Newton's inverse-square law

of universal gravitation. The heavily tested law can be used to derive Gauss's law and vice versa.

The phenomenon of static electricity had been observed and investigated long before Coulomb, however. Ancient Mediterranean cultures knew that certain objects could be rubbed with an animal's fur until they attracted lighter objects. Around 600 B.C.E., Thales of Miletus made observations on static electricity which led him to believe that friction rendered amber magnetic, while minerals such as magnetite needed no rubbing. Unfortunately, Thales was incorrect in attributing the attraction to magnetic effect, but science did later prove a link between electricity and magnetism.

Until 1600, electricity would remain little more than an intellectual curiosity, at which time William Gilbert studied magnetism and electricity and distinguished static electricity (which was produced by rubbing amber) from the lodestone effect. From this, he coined the New Latin word *electricus* (like amber or of amber, from the Greek word *elektron*, which also meant amber). Today, we have the English words *electric* and *electricity*, both of which were printed in Thomas Browne's *Pseudodoxia Epidemica* of 1646.

Several scientists in the early 18th century, including Alessandro Volta and Daniel Bernoulli, suspected that electrical force diminished with distance in the same way that the force of gravity did. These two men measured the force between the plates of a capacitor.

In 1758, Franz Aepinus proposed the *inverse-square law*. Based on this law and experiments with electrically charged spheres, Englishman Joseph Priestley was among the first to propose that electrical force followed this law—similar to Newton's law of universal gravitation, though he did not elaborate on this theory. In 1767, he argued that the force between charges varied as the inverse square of the distance.

Coulomb's Torsion Balance

Continuing with Coulomb's original work, Scottish physicist John Robinson announced that the force of repulsion of two positive or two negative circles varied as $x - 2.06$, thereby solving an age-old question. By the early 1770s,

dependence on the force between charged bodies upon charge and distance had been discovered (yet not published) by the Englishman Henry Cavendish.

In 1785, Coulomb published his first three reports on magnetism and electricity and stated his law. This led to the development of the field of magnetism. Coulomb utilized a torsion balance to study the attraction forces and repulsion of charged particles. He determined that the magnitude of the electric force between two point charges is proportional to the square of the distance between them.

This torsion balance is composed of a bar suspended from its center by a thin fiber, which acts as a very weak torsion spring. When Coulomb performed his famous experiment, he used an insulating rod as the torsion balance, along with a metal-coated ball attached to one end that was suspended by a silk thread. This ball was charged with static electricity, while a second ball with the same charged polarity was brought close to it. The two balls repelled one another, and the fiber was twisted through a certain angle, which could be read on the instrument's scale. Coulomb was then able to calculate the force between the balls and determine his inverse-square proportionality law by knowing how much force it took to twist the fiber through a given angle.

Coulomb's Law

According to Coulomb's law, the magnitude of the electrostatic force of interaction between two point charges is exactly proportionate to the scalar multiplication of the magnitudes of charges. He also asserts that is also inversely proportional to the square of the distance between these two points. The force along the straight line joins the points, and if their two charges have the same sign, the electrostatic force between them will be repulsive, whereas if they have different signs, the force between them is attractive.

Coulomb's law can be stated as a mathematical expression as follows:

$$F_1 = k_e \frac{q_1 q_2}{|r_{21}|^2} \hat{r}_{21} \text{ and } |F_1| = k_e \frac{|q_1 q_2|}{r^2},$$

respectively, where k_e is Coulomb's constant,

$$(k_e = 8.9875517873681764 \times 10^9 \text{ N} \cdot \text{m}^2 \cdot \text{C}^{-2})$$

q_1 and q_2 are the signed magnitudes of the charges, the scalar r is the distance between the charges, the vector $\mathbf{r}_{21} = \mathbf{r}_1 - \mathbf{r}_2$ is the vectorial distance between the charges, and $\hat{\mathbf{r}}_{21} = \mathbf{r}_{21}/|\mathbf{r}_{21}|$ (a unit vector pointing from q_2 to q_1). The vector form of the equation calculates the force $\mathbf{F}_1$ applied on q_1 by q_2 . If $\mathbf{r}_{12}$ is used instead, then the effect on q_2 can be found. This can also be determined by employing Newton's third law: $\mathbf{F}_2 = -\mathbf{F}_1$.

Units

Electromagnetic theory is typically expressed using the standard International System of Units (SI units). Charge is then measured in coulombs, force in newtons, and distance in meters. Coulomb's constant is represented as $k_e = 1/(4\pi\epsilon_0\epsilon)$. The constant ϵ_0 is the permittivity of free space in $\text{C}^2 \text{m}^{-2} \text{N}^{-1}$. And ϵ is the relative permittivity of the material in which the charges are immersed. It is also dimensionless.

Coulomb's law and Coulomb's constant can be interpreted in various terms: One of these is electronic units, also known as *Gaussian units*. In both sets of units, the unit charge (statcoulomb or esu) is defined in a way that leaves the Coulomb constant k to disappear, owing to the fact that it has a value of one and becomes dimensionless.

The other unit is atomic units. In these units, force is expressed in hartrees per Bohr radius, the distances in terms of the Bohr radius and the charge in terms of elementary charge.

There is much more to say about Coulomb's law in terms of its theory and practical applications that space limitations will not enable us to discuss here, including the areas of continuous charge distribution, the system of discrete charges, vector form, scalar form, the electric field, Coulomb's relationship to atoms, and Coulomb's constant. It is most important, however, to note that there are *two conditions for the validity* of Coulomb's law: The charges considered must be point charges, and they should be stationary with respect to one another.

The Enlightenment, Science, and Religion

Physics, referred to early on simply as *science* in many cases, had a large impact on religious views

during the Enlightenment. We have already discussed Priestley's religious views and how they coincided with the Age of Enlightenment, but now let us look at how they played out in society in general.

Much in the religious change of thought revolved around the thoughts of thinkers and theorists such as Newton and John Locke. Their ideas highlighted the transformation of society and of describing knowledge in terms of human experience instead of from a Biblical perspective.

John Locke

Western Europe was rapidly turning to urbanization, due to the wealth coming in from colonialism. Many people left the farms for the cities, and the close proximity of populations allowed thinkers and scholars to meet and share ideas. Large, prominent cities such as Edinburgh, Paris, and London became known as strongholds of Enlightenment thought. Following the decline of Catholicism, English culture flourished and began to produce some prominent engineers, philosophers, and scientists as wealth flowed in from mercantile and colonial ventures in Asia and the New World.

Scholars adopted empiricism during this Age of Reason, and they believed that all theories should be based upon human experience and personal observations. The hand of God was no longer seen behind every phenomenon, and science began to take precedence. Some scholars, like Newton, did feel that there was room for a divine creator, the Uncaused Cause of Aristotle. The world's new definition of knowledge was embedded in every aspect of society, and the rapid accumulation of knowledge led to science being divided into separate disciplines.

Philosophers and scholars rebelled against the restrictions of Christianity and instead used metaphysics and science to question the universe. Europe was becoming secular, and the field of science itself moved away from Biblical absolutism and realism. Philosophers such as Descartes had previously questioned the nature of the soul in the 1600s and had envisioned a universe that was purely mechanical and physical in nature. Descartes believed that the human body and animals were automatons, with the soul alone elevating humanity.

Financial support was found for new research on these topics, and new inventions such as the barometer, telescope, and microscope allowed scholars to make more accurate observations than ever before, while also conducting better experiments as the scientific method was refined. The printing press had made books much cheaper, and roads and improved transportation allowed for the flow of traffic, as well as ideas. Intellectuals such as Leibniz and Newton conducted fierce debates by letter. Scientific societies also provided a place to share and refine ideas, as well as receive constructive criticism.

Conclusion: The Enlightenment and Social Reform

Alongside amazing scientific discoveries, much progress was made in social reform, owing to the work of Enlightenment thinkers. They provided challenges to theocracy and autocracy, which had been fundamental elements of society for many centuries. Commerce and trade vied with agriculture for importance, and Europe became transformed into a host of societies that increasingly pursued the pleasures of life.

The growth of capital also allowed scholars the resources they needed and led to a sense of optimism—particularly, a feeling that humanity could change the world for the better. Philosophers began to try to understand the nature of knowledge itself instead of focusing on abstract musings about the universe and Aristotelian metaphysics. They began to realize that humankind had the capacity to influence nature, and not everything was necessarily foreordained by divine beings and supernatural forces. More and more, humanity was considered to be governed by natural law. Gradually, the natural sciences—chemistry, astronomy, and physics—gave rise to the political and social sciences, as scholars came to believe that the world, including human society, could be understood through methodical observation and study.

Katie Moss

See also Philosophy of Science; Physics, 19th Century; Physics, 20th Century; Physics, Antiquity; Physics, Contemporary; Physics, Middle Ages; Physics, Renaissance

Further Readings

- American Chemical Society International Historic Chemical Landmarks. (2018, August). *Discovery of oxygen by Joseph Priestley*. <http://www.acs.org/content/acs/en/education/whatischemistry/landmarks/josephpriestleyoxygen.html>
- Best, N. W. (2015). Lavoisier's Reflections on phlogiston I: Against phlogiston theory. *Foundations of Chemistry*, 17(2), 361–378.
- Brown, P. L. (1985). *Halley and his comet*. Blandford Press.
- Franklin, B. (1791). *The autobiography of Benjamin Franklin*. Various reprinted editions.
- James, I. M. (2004). *Remarkable physicists: From Galileo to Yukawa*. Cambridge University Press.
- Kramnick, I. (Ed.). (1995). *The portable enlightenment reader*. Penguin.

PHYSICS, GRAVITATIONAL THEORY

A gravitational theory explains, first of all, why heavy objects fall to the ground; Isaac Newton's (1643–1727) epochal formulation of gravitational theory in the 1680s extended gravity universally on Earth and into the heavens, the realm of astronomy; his laws of motion and his theories in other areas of physics would hold sway for more than two centuries. Subsequently, the move from Newton's *action at a distance* celestial mechanics to local field theories made it appropriate, especially once Albert Einstein's special theory of relativity had emerged in 1905, to seek a relativistic field theory of gravity. A physical field associates a number (or more) to each point in space at a time; hydrodynamics is a familiar example, but most fields are less tangible.

Whereas the first satisfactory theory of gravity of this type was Gunnar Nordström's, in the same years (early 1910s) Albert Einstein (1879–1955) was seeking a generalized theory of relativity to include gravitation. Einstein's 1915–1916 theory, which has come to be called *general relativity*, was revolutionary not only in dethroning the long-reigning Newtonian physics but also in advancing the idea that space-time is curved, indeed variably curved, a claim that only made sense given recent mathematical advances unknown to most

physicists at the time and bewildering to the Euclidean geometry-trained common person. With the bending of starlight by a strong gravitational field, observed by astronomer Arthur S. Eddington in 1919, the seemingly incomprehensible was empirically verified and Einstein took on a prophetic stature in the public imagination. This entry discusses the language and content of Einstein's theory, its relation to experiments, its relation to cosmology, conceptual issues involving the century-long controversy about gravitational energy and conservation laws and a recurring debate about determinism, the question of gravitational exceptionalism versus egalitarianism vis-à-vis other physical fields, and how gravitational theory might be merged with quantum mechanics into quantum gravity.

Language and Content of Einstein's Theory

Einstein's field equations are 10 coupled nonlinear partial differential equations for the space-time metric tensor, which relates small space-time coordinate differences to physical distances and times. (In geography, relating longitude to distance requires the metric tensor.) These partial differential equations relate the metric and its first and second partial derivatives (rates of change with respect to coordinates) to nongravitational energy, momentum densities, and stress or pressure. The equations display some structural similarity to James Clerk Maxwell's (1831–1879) field equations for electromagnetism, but Einstein's equations are nonlinear: The sum of two solutions is not a solution. Both theories admit waves propagating at the speed of light. A novel feature of general relativity is the infinitely greater variety of equally appropriate ways to give numerical labels (coordinates) to times and places. This *general covariance*, the preservation of the form of the laws under basically arbitrary coordinate transformations, requires tensor calculus. Traditional tensor calculus involves a taxonomy of the various ways that physical quantities can transform under change of a coordinate system. A tensor of a certain sort has values relative to a coordinate system, related to values in any other coordinate system by a transformation rule involving partial derivatives. Especially among mathematicians there has been an effort to reformulate tensor

calculus and differential geometry (the calculus-based geometry of curved spaces) using coordinates, but rarely. In key instances, different sets of components of a tensor appear as facets of an invariant abstract entity. This entity is analogous to a Platonic proposition expressed by sentences such as “Snow is white” and (in German) “Schnee ist weiss.” To what degree this *coordinate-free* approach retains classical expressive power seems unclear presently, but the approach is advantageous for global problems.

Already in Newtonian physics, gravity is universal: A point-sized bit of matter experiences (and causes) a gravitational force proportional to its mass—the same mass that appears in Newton's second law determining acceleration. Equality of active and passive gravitational mass with inertial mass suggests a tight connection with space and time. Experimentally, this result could be approximate. Einstein took it as exact in his theory of general relativity, merging gravity into space-time geometry. Hence, one cannot speak objectively of the gravitational force on an object. Particles can be viewed as moving on straight lines (*geodesics*) in a curved space-time.

Einstein took a key feature of his theory to be the lack of preferred inertial coordinate systems, a feature he called *general covariance*. Among physicists this claim has been accepted as basically right, but it has resisted clarification despite decades of effort; the newer term *background independence* covers similar ground. General relativistic space-time geometry generally lacks symmetries—a claim not to be confused with the question of the symmetry of laws, which general relativity has in novel abundance. For earlier theories, the two questions hardly need distinguishing, but for general relativity, conflation has serious consequences for conservation laws.

Gravitational Theory and Experiment

Early observations favoring Einstein's theory over rival theories in the 1910s included its better treatment of the perihelion precession of Mercury (an issue complicated by uncertainties about the matter content of the solar system) and, strikingly, the bending of light in 1919—a phenomenon not readily accommodated by Newton's theory and clearly not predicted by Nordström's theory. The

1950s experiment by Robert V. Pound and Glen A. Rebka, Jr. observed the gravitational red/blueshifting of light. In 1981 Russell A. Hulse and Joseph H. Taylor, Jr. found that a binary pulsar orbit was decaying according to the gravitational radiation predicted by Einstein's equations, in contrast to most other theories of gravity. Gravitational time dilation is crucially used in global positioning systems.

In local field theories with laws having at least the symmetries characteristic of special relativity (including general relativity), waves travel at light speed. Thus it would take a billion years to see or feel (gravitationally) an event a billion light years away. Cosmic expansion, according to Big Bang cosmology, complicates that result slightly. A conceptual wrinkle arises, owing to general relativity's radical relativity of simultaneity: One could, for example, define simultaneity using a past-facing hyperboloid instead of a roughly flat hypersurface, and so regard events 10 billion light years away and 10 billion years earlier (by naïve definitions) as not very far in the past, if one wishes. Hyperboloidal time coordinates are useful for finding approximate solutions of Einstein's equations using computers. In the 2015–2016, after decades of effort in the laser interferometer gravitational-wave observatory (LIGO) program (among others), gravitational radiation was detected. Only large, fast-moving masses produce appreciable radiation, such as binary systems involving neutron stars or black holes.

Gravitational attraction so strong that even light cannot escape is predicted by the theory of general relativity in some contexts. In the 1960s, the tendency of a sufficient quantity of localized matter to collapse toward a singular point because of gravity was found to be generic, using new global techniques employed by Roger Penrose and Stephen Hawking. Gravitational collapse tends to hide itself behind a horizon, an outermost surface on which *outbound* light remains at fixed radial distance. It is therefore difficult to know much about matter collapsed into a black hole. Whereas experimental particle physics can largely ignore gravity, efforts beginning in the 1970s to embed particle physicists' quantum field theory in a gravitational environment has encouraged an analogy (or more) between black holes and thermal physics.

Cosmology

The typical cosmological model describes space as homogeneous and isotropic (the same in all places and directions), an assumption motivated by pragmatism (tractable mathematics), empiricism (astronomical data fit tolerably well), and naturalistic philosophy (denying a *special* world). Attempting to express tensor calculus verbally, one can say that the universe is expanding, or that everything is receding from everything else. This expansion redshifts light, including the cosmic microwave background and light from stars, increasingly with distance. Extrapolating backwards and trusting Einstein's equations beyond their expected realm of validity (for one would expect quantum gravity effects to make an important difference) implies moments of arbitrarily great density and temperature. The Big Bang singularity “at” $t = 0$, however, is not a consensus scientific explanation of the origin of the universe; experts typically expect quantum gravity to smooth out such a backward extrapolation. This debate has little theological relevance: While the Big Bang does not support a creation event, quantum gravity's removing the singular Bang from the model wouldn't exclude creation from reality either. Claims of cosmic fine tuning for life continue to develop and be refined; some might be alternatively explained by way of a multiverse.

Big Bang nucleosynthesis uses nuclear physics, with the universe assumed near equilibrium in a hot expanding phase, to predict the proportions of light element isotopes (hydrogen, helium, lithium, etc.). Leaving heavier elements to be predicted by nuclear astrophysics including supernovae, Big Bang nucleosynthesis predicts light element abundances rather well at the time of writing.

Although an expanding universe would be expected to decelerate because of gravitational attraction, several puzzles, including how flat space is, could be resolved by positing vastly more rapid expansion early on. This *inflation* admits various moderately plausible theoretical explanations in terms of one or more fields, dubbed the *inflaton*, which, however, typically lack independent support. Inflaton behavior constitutes an effective positive *cosmological constant* Λ (capital Greek letter lambda), a modification to his field equations that Einstein suggested in 1917.

Since the late 1990s, a combination of two large-distance empirical puzzles for general relativity has opened the field theoretically to explore other theories. Since the mid-20th century, an awkward find for general relativity at large distances has been *dark matter*, the inferred cause of the observed failure of galaxies to disintegrate, despite the insufficiency of visible matter to bind them. The name *dark matter* implies that the gravitational laws typically were not blamed. Remarkably, in 1998, observations of supernovas indicated that the universe is currently in a state of accelerating expansion ($\Lambda > 0$). (The alternative proposal of inhomogeneity has not fared well.) This phenomenon (or its explanation) is *dark energy*. Whether dark energy is gravitational or material is unclear. The combination of dark matter and dark energy, both arguably flawed predictions on large distances, broke the nearly universal consensus on general relativity. Though Einstein's theory remains favored, there is increased work on alternative theories.

Interpretive Questions

In post-1910s physics, the conservation of energy and momentum are derived from the uniformity of natural laws over time (energy) and place (momentum). Formally, the same treatment works for general relativity, but the result is a *pseudotensor*: The descriptions of energy and momentum in different coordinate systems appear inequivalent. Even the absence or presence of gravitational energy at a point is not invariant. One appears to have a choice of saying that gravitational energy exists but is not objectively at specific space-time points (*nonlocalizability*, the mainstream view), or does not exist despite the formal conservation of material energy + gravitational pseudo-energy in any coordinate system, or is infinitely multiple (a 2000s idea). The popular *quasi-local* program aims to localize energy to finite regions.

Einstein's process of discovery included the *hole argument*, by which he temporarily convinced himself that generally covariant theories were indeterministic: One past has multiple possible futures. This concern was resolved by way of the *point-coincidence argument*, which individuates space-time points by their material contents (somewhat as in Gottfried Wilhelm Leibniz's

thought from the 1710s), not primitively as the hole argument assumed (akin to Newton's ideas at times). For reasons likely involving coordinate-free differential geometry as well as general relativity's historiography, a revived hole argument appeared in philosophy in the 1980s and remains widely discussed. While no one denies determinism (locally) in Einstein's theory, ways to make space-time points' identities depend on physical properties vary.

A longstanding divide concerning gravitation is whether gravity is exceptional or is basically like the other fundamental interactions (electromagnetism and the weak and strong nuclear forces). The former view is usual among specialists in general relativity; the latter egalitarian view is usual among particle physicists, including luminaries such as Richard Feynman and Steven Weinberg. One's view on this issue affects many other questions. First, does one expect dramatic global consequences and hence need global techniques including coordinate-free geometry? Second, what modifications or alternative theories are plausible—theories generalizing metric geometry or generalizing a theory of a field/particle that exactly fits a particular slot in particle physicists' taxonomy? Third, how should one merge gravity with quantum mechanics into quantum gravity? General relativists have preferred quantization techniques admitting dramatic or global novelty, whereas particle physicists have tended to assume small perturbations. Fourth, how does one approach conceptual puzzles about general relativity—for example, about energy (non)conservation, or the apparent absence of change in some quantum gravity programs—are they insights or, at best, options? Fifth, what does one consider relevant, insightful, accidental, important, or mistaken in the history of general relativity? In physics this divide started weakening in the 1970s with the advent of supergravity and later (super)string theory. Supergravity ascribes to gravity a *gravitino* partner that is fermionic, satisfying the Pauli exclusion principle.

Since the 1990s, historians of physics have made an exhaustive study of Einstein's process of discovery, including unpublished notebooks. A key result is the importance of a second line of thought in Einstein's work, which he retrospectively suppressed: Besides his mathematical

strategy of principles such as general covariance, generalized relativity of motion, and the equivalence of gravity and acceleration, Einstein employed a physical strategy including conservation laws and analogies to Newtonian gravity and to electromagnetism. Einstein retrospectively credited the mathematical strategy with success and further pursued into theories aiming to unify electromagnetism into the same framework, but this work is generally considered unsuccessful. It turns out that particle physicists' rigorous rederivation of Einstein's equations (c. 1938–c. 1972) uses the same kinds of ideas as Einstein's physical strategy. Hence, whether or not Einstein's physical strategy succeeded in his hands, it does work.

Quantum Gravity

All the other fundamental interactions are described by quantum field theories. Presumably gravity also must be *quantized*, perhaps on pain of contradiction. While guidance from experiment would be ideal, that prospect seems remote at the time of writing. Whether gravity is exceptional or just another force strongly influences one's views about quantization.

Gravity as the great exception is likely to yield dramatic, novel, fundamentally nonlinear phenomena requiring Hamiltonian (*canonical*) quantization or path integration. Loop quantum gravity continues this tradition effectively. A puzzle of apparently losing time and change has accompanied canonical quantum gravity since the 1950s: It has been difficult to identify the physically real content, which tends to get submerged in the conventional. A reforming viewpoint since the 1980s suggests recovering early 1950s mathematically rigorous equivalence to the four-dimensional symmetry of Einstein's equations.

Gravity as just another fundamental force is likely yield modest perturbative effects. This project saw considerable effort from particle physics in the 1960s and 1970s but produced incurable infinities (*nonrenormalizability*). This problem might be resolved if there is a special fermionic partner(s), called a *gravitino*, of the gravitational particle (graviton) that introduces minus signs so as to cancel bad terms from gravity, as in 1970s+ supergravity. The project further developed into

(super)string theory. Any such *supersymmetric* particles remain elusive empirically, however. A 1990s+ shift to *effective* field theories indicates that the old infinities are not fatal if new physics might become relevant at sufficiently high energies. On that view, we can know enough about gravity without knowing everything.

J. Brian Pitts

See also Cosmology; Laws of Nature; Laws (Scientific); Physical Theory; Physics, 20th Century

Further Readings

- Blum, A. S., & Rickles, D., (Eds.). (2018). *Quantum gravity in the first half of the twentieth century: A sourcebook*. Edition Open Access, Max Planck Institute for the History of Science.
- Feynman, R. P., Morinigo, F. B., Wagner, W. G., Hatfield, B., Preskill, J., & Thorne, K. S. (1995). *Feynman lectures on gravitation*. Addison-Wesley. (Originally by California Institute of Technology, 1963)
- Howard, D. (1999). Point coincidences and pointer coincidences: Einstein on the invariant content of space-time theories. In H. Goenner, J. Renn, J. Ritter, & T. Sauer (Eds.), *The expanding worlds of general relativity* (pp. 463–500). Birkhäuser.
- Kennefick, D. (2007). *Traveling at the speed of thought: Einstein and the quest for gravitational waves*. Princeton University Press.
- Norton, J. D. (1993). General covariance and the foundations of general relativity: Eight decades of dispute. *Reports on Progress in Physics*, 56, 791–858.
- Ohanian, H., & Ruffini, R. (2013). *Gravitation and spacetime* (3rd ed.). Cambridge University Press.
- Renn, J. (Ed.). *The genesis of general relativity, volume 1: Einstein's Zurich notebook: Introduction and source*. Springer.

PHYSICS, MIDDLE AGES

The Middle Ages, also referred to as the *medieval period*, took place from approximately the 5th to the 15th century C.E. It began with the gradual collapse of the Western Roman Empire and was followed by what is now known as the *Renaissance and the Age of Discovery*. It is the second of

three major stages of Western history: antiquity, the medieval period, and the modern period. The medieval period itself is divided by historians into the Early, High, and Late Middle Ages. This entry describes the origins of medieval physical theories and practices in the work of earlier scholars and the efforts made during the Middle Ages to preserve and build upon ancient knowledge. Prominent thinkers of the period in Europe and in the Middle East and their notable achievements are discussed. A section on the growth of trade throughout and beyond Europe and the founding of centers of learning is followed by an exploration of the scientific and mathematical culture of the High and Late Middle Ages, including a description of the work of scholars at Merton College, Oxford, in helping to establish the foundations of the scientific method.

Philosophers/Physicists Influential on Thought in the Middle Ages

The Middle Ages inherited a large amount of knowledge from the age of antiquity, including the philosophies of Aristotle and other Greek thinkers, and work in ancient astronomy, owing to the efforts of the polymathic scholar Ptolemy, who lived in the 2nd century C.E. Aristotle, like philosophers before him, stated that terrestrial and celestial realms were separated by the orbit of the Moon and only differentiated themselves by their unique physical behaviors. Aristotle's terrestrial physics ideas were founded on the existence of four elements (fire, water, air, and earth) and the assertion that every motion requires a specification of cause for its motion.

Aristotle believed that there were two main types of motion: violent motion (when an object is removed forcibly from its natural place) and natural motion (the motion arising from the nature of an object). This model dictated that the natural place of the element earth was at the center of the cosmos and that Earth, the human domain, was the center of the heavens.

Although the terrestrial realm featured constant change, the heavenly bodies moved in their circular, uniform orbits and were unchanging. Ptolemy, drawing on a massive amount of astronomical data, modeled the orbit of each heavenly body with a system of circular motions. He was

often forced by new data to make additions to his model, such as his *eccentric* model and his *equant* model.

Although this work was enormously valuable, it was lost to the West over the course of the decline and dissolution of the Roman Empire. A number of Islamic scholars, however, were interested in it and translated it into Arabic, where they studied it extensively, eventually making their own major innovations in matter theory, optics, astronomy, and mathematics (including what we still refer to as *Arabic numerals*).

Around the 8th century, the Syrian astronomer al-Battani improved upon Ptolemy's orbits of the Moon and Sun, created a revised catalogue of stars, and constructed astronomical instruments. Avempace (an Arab Andalusian also known as *Ibn Badja*) developed a philosophy first outlined by Neoplatonist philosopher John Philoponus in the 6th century C.E. This argument claimed that the amount of time it took for an object to fall was proportional to its weight. After the Muslim conquest of Hispania during the 8th century, ancient knowledge came to the Latin west once again. Arab scholars such as Averroes (also known as *Ibn Rushd*) became interpreters of Aristotle—providing translations that were much closer to the original texts than they previously had been.

Early Medieval Society: The Dark Ages

The early medieval period lasted from about 500 to 1000 C.E. and is commonly regarded as the *Dark Ages* because of the barbarism and ignorance that had become prevalent across Europe. Yet even this era saw advances in science and technology; unfortunately, most of this information was lost due to constant warfare on the European continent.

It was monastic study that kept literacy and learning alive. Although most monasteries centered their pursuits on the Bible and early Christian theologians, the monks also studied astronomy, geometry, medicine, and mathematics. In the secular realm, tradesmen also made advances in the fields of science and mechanics. It is true, however, that little progress was made in terms of the scientific method. In Europe, classical thought and philosophy were mostly forgotten—though they were preserved in Islam and Byzantium.

Medieval Thoughts on Space

Following in the path of Aristotle, medieval scholars believed that the celestial realm was one of unchanging perfection and that the Moon, the planets, the Sun, and the sphere of the fixed stars rotated around the Earth in their own celestial spheres. Ptolemy added epicycles in addition to Aristotle's concentric spheres, which led to medieval astronomers calling heavenly bodies either *total orbs* or *partial orbs*.

These orbs, it was believed, could communicate rotational movement to one another with no resistive force and were composed of quintessence or ether, which was conceived as a transparent, ageless substance. This is where the Christian God came into play. Taking the place of Aristotle's *Unmoved Mover*, he was believed to be beyond the outermost sphere of the fixed stars and to have the ability to impress an impetus on each orb at the moment of creation. Because there had been no resistance put upon them since that time, they were still rotating. Nicole Oresme and Jean Buridan, two 14th-century French scholars, considered the possibility of using a rotating Earth to explain the diurnal motion of the fixed stars, but their arguments did not go very far, as they themselves remained unconvinced.

The Middle East

After the Roman Empire was divided into Eastern and Western halves by the emperor Constantine in the 4th century C.E., the Western Roman Empire lost much of its history. In the Middle East, Greek philosophy was taken up by some scholars in the newly created Arab Empire. A period of Muslim scholarship arose with the spread of the Islamic religion in the 7th and 8th centuries, and this Islamic Golden Age lasted until the 13th century. The scholarship was furthered by the use of a single language, Arabic, as well as access to Greek texts from the Byzantine Empire and Indian sources of scholarship.

The Eastern, or Byzantine, Empire housed centers of learning such as Constantinople, while Western Europe's knowledge was preserved and cultivated in Christian monasteries until medieval universities were developed in the 12th and 13th centuries. Monastic schools taught ancient texts

and new works on practical subjects such as time-keeping and medicine.

Muslim scientists placed emphasis on experimentation, leading to an early scientific method being developed in the Muslim world and a considerable progress being made in methodology. Al-Haytham (Alhazen) began experiments in optics in ca. 1000 in his *Book of Optics*. He is known as the founder of optics because of his empirical proof of the intromission theory of light. He has even been called the *first scientist* for his work in developing what we now call the scientific method.

The most important use of this scientific method by Muslim scientists was to host experiments to distinguish between competing scientific theories comprised within a generally empirical orientation. In the field of mathematics, Persian mathematician Muhammad ibn Musa al-Khwarizmi created the concept of the algorithm (and donated his name to the field), while the term *algebra* comes from *al-jabr*, the leading title of one of his publications. What we commonly refer to as Arabic numerals originated from India, but Muslim mathematicians gradually refined the number system by introducing such edits as decimal point notation.

Persian scholar Al-Razi contributed to medicine and chemistry, and Sabian mathematician Al-Battani made excellent contributions to the fields of mathematics and astronomy. Al-Battani improved the measurements of Hipparchus, as they were kept safe in the translation of Ptolemy's *Hè Megalè Syntaxis* (The Great Treatise), translated as *Almagest*. He also made improvements to the measurement of the precession of the Earth's axis. The corrections made to the geocentric model by al-Battani, Ibn al-Haytham, Averroes, and the Maragha astronomers such as Nasir al-Din al-Tusi, Ibn al-Shatir, and Mo'ayyeduddin Urdu are similar to the heliocentric model later introduced by Nicolaus Copernicus in 1543. Heliocentric theories may have also been discussed by several other Muslim astronomers such as Ja'far ibn Muhammad, Abu Ma'shar al-Balkhi, Najm al-Din al-Qazwini al-Katibi, Abu-Rayhan Biruni, Quth al-Din al-Shirazi, and Abu Said al-Sijzi.

In the field of chemistry, Muslim chemists and alchemists play an important role, and scholar Jābir ibn Hayyān is considered by many to be the father of chemistry. His works later influenced

Roger Bacon, who introduced the empirical method to Europe.

There were many other famous Islamic scientists from this period, all of which will not be named, though we will mention Ibn Sina (Avicenna), who is thought to be the region's most influential philosopher, as he pioneered the science of experimental medicine and became the first physician to conduct clinical trials.

Islamic science began to decline by the 12th and 13th centuries, due to the Mongol conquests, which destroyed many universities, libraries, and hospitals. The end of the Islamic Golden Age is marked by the destruction of the intellectual center of Baghdad, the capital of the Abbasid caliphate in 1258.

The Expansion of European Exploration and Thought

During the 9th century, Western Europeans came to recognize that systemized education was the key to prosperity and unity. In his period, known as the Carolingian Renaissance, the Frankish king Charlemagne (Charles the Great, crowned as Emperor of the Romans in 800 C.E. by Pope Leo III) sought to establish knowledge as an integral part of medieval society. He may have been brutal in war, but he was a great believer in education, art, and culture, and he used the Catholic Church to transmit this information to the public. Although he was himself illiterate, he ordered the translation of many Latin texts and promoted the study of astronomy.

In England, a monk called Alcuin of York (ca. 724–804) established a system of education in theology, art, geometry, arithmetic and astronomy. He also promoted the establishment of schools. However, his and Charlemagne's goals would be ended by Charlemagne's death and the end of the Frankish Empire (and the beginning of barbarian incursions). However, some centers of learning clung tightly to the pursuit of scholarship, giving rise to a number of extraordinary thinkers in the years that followed.

The High Middle Ages

The period from 1000 to 1300 is referred to by historians as the High Middle Ages, when the long cycle of wars in much of Western Europe began to

diminish. Populations grew, and Christianity provided a unifying force and a sense of shared identity across the continent.

Sharing of ideas (promoted through trade) was common, and ideas were brought to Western Europe from the Arabic empires. Because the Muslims had translated many ancient Greek texts into Arabic, scholars from around Europe were able to translate them back from Arabic into Latin. This way, Greco-Roman knowledge passed through Europe, where places of academic study (*studium generale*) were becoming universities. Many scholars, such as Gerard of Cremona (c. 1114–1187) learned Arabic so they could complete these translations.

In the 12th century, *studium generale* were emerging across Western Europe, combining the knowledge of the ancient Greeks with new discoveries by Muslim scientists and philosophers. They drew scholars from near and far, and their unique curriculum inspired the basis of Christian scholasticism. Although most of the study turned toward theology, the schools also began to integrate scientific empiricism.

Great thinkers of this period included such men as Robert Grosseteste, Thomas Aquinas, Roger Bacon, and William of Ockham—who all worked diligently to perfect and refine what we know today as the scientific method.

Thomas Aquinas was extremely interested in using philosophy to prove the existence of God; he also oversaw a change from Platonic reasoning toward Aristotelian empiricism.

Oxford's Francis School was founded by Robert Grosseteste, who was one of the major contributors to the scientific method. He began to promote the dualistic scientific method as proposed by Aristotle. Grosseteste had an idea that composition and resolution involved prediction and experimentation. He believed that observations could be utilized to proposed universal laws and predict outcomes. Grosseteste was one of the first scientists to use this method as an empirical process, and his ideas influenced Galileo and filtered into the 17th-century Age of Enlightenment. However, the greatest impact on the scientific method was made by Grosseteste's pupil, Roger Bacon.

Roger Bacon was one of the greatest minds of the medieval period and a refiner of the scientific method. He took the works of Aristotle, the Islamic

alchemists, and Grosseteste and used them to propose the idea of induction as the key element of empiricism. He described the method of observation, the hypothesis, and experimentation, noting that results should always be independently verified. He documented all notes in detail so that his experiments could be repeated. Today, any student who writes an experimental paper must use this method.

The Late Middle Ages

By the time of the late Middle Ages (1300–1500), a revival of science and learning in the West was under way. A 13th-century mathematician by the name of Jordanus de Nemore had pioneered a series of studies revolving around static bodies. Along with study levers, he also analyzed the (lower) apparent weight of a mass that was static on an inclined plane.

William of Ockham, in the 14th century, proposed his idea of parsimony and the famous principle of *Occam's razor*, which is still used by scientists today to determine answers between conflicting explanations. Around the same time, Jean Buridan developed the idea of impetus, a concept that predated Newtonian physics and inertia (and challenged Aristotelian physics).

Despite the fact that the church condemned several radical interpretations of Aristotle's work during the late 13th century, there was a flurry of activity around these activities during the next century, particularly revolving around the subject of motion. Merton College at Oxford and the University of Paris became centers of activity for this research. Prominent mathematicians included John Buridan and Nicole Oresme in Paris and Richard Swineshead and Thomas Bradwardine in England.

Merton College

In England, academicians at Merton College, Oxford, presented a distinct separation between *dynamics*, in which the causes of motion are specified, and *kinetics*, in which motion is described. They examined the dynamical problems occurring in the context of Aristotelian physics, such as the problem of projectile motion.

Many other advancements were made in kinematics, which is the release from the search for causation. The Mertonians developed the concept

of velocity in analogy with the idea of the intensity of a quality (such as the greenness of grass). They also made sense of the difference between nonuniform (accelerated) and uniform (constant velocity) motion. The Mertonians also introduced the first instance of the *mean velocity theorem*, which gave the ability to compare uniform motion to constant-acceleration motion.

Mertonians explained their analyses of motion through the use of words, but other scholars used graphical techniques to explain their theories. Perhaps the most impressive presentation of the mean speed theorem was given by Oresme. He recorded the duration of motion along a subject line and showed the intensity of the velocity as a sequence of varying vertical line segments. Although his work was predominantly abstract (rather than based on experimentation), it helped to inspire Galileo's later work in kinematics.

Katie Moss

See also Physical Theory; Physics, 19th Century; Physics, 20th Century; Physics, Contemporary; Physics, Enlightenment; Physics, Renaissance

Further Readings

- Dorminey, B. (2020, October). Tech Lessons from the "Dark Ages." *Forbes*. <http://www.forbes.com/sites/brucedorminey/2014/08/30/tech-lessons-from-the-dark-ages/>
- Hannam, J. (2011). *The genesis of science: How the Christian Middle Ages launched the scientific revolution*. Regnery Publishing.
- Lindberg, D. C., & Shank, M. H. (2013). *Cambridge history of science: Medieval science* (Vol. 2). Cambridge University Press.
- Shuttleworth, M. (2021, October). Middle-Ages science. *Explorable*. <https://explorable.com/middle-ages-science>

PHYSICS, 19TH CENTURY

The 19th century was an era of rapid growth and transformation in physics, with a quest for the mathematical laws behind observed phenomena and an increased emphasis on quantification. The

establishment of well-funded institutes housing laboratories for conducting physical research both for private concerns and for the state gave rise to sophisticated technologies and the elicitation, study, and measurement of a wide range of phenomena. Foremost, perhaps, among scientists' goals was the creation of a unified physics—the formulation of the fundamental laws that govern all physical phenomena. Physics also catered to the needs and imaginations of different fields—political, military, and commercial—within the expanding and burgeoning industrial nations. During this period, the so-called imponderables (magnetism, electricity, heat, and light) were the main objects of study. As the century advanced, these were treated as having originated from the same fundamental source of energy but with different manifestation. At the same time, the ancient notion of the ether, a substance supposed to pervade all space beyond the sphere of the moon and to transmit light and heat, came into question.

This entry focuses on several areas of advancement in physics during the 19th century, including magnetism and electricity, heat, radiation, thermodynamics, elasticity, dynamics, hydrodynamics, and optics. Other domains of physics in which major developments occurred during this period include crystallography, capillarity, diffusion, viscosity, aeromechanics, acoustics, kinetic theory of gases, velocity of light, photochemistry, fluorescence, and photometry. Owing to space limitations, a detailed account of the latter topics is beyond the scope of this entry. Readers are encouraged to consult the bibliography below for further information.

Key Developments

The study of electric currents received a major boost with the invention of the *Voltaic pile*, the electric battery created by Alessandro Volta in 1800. Shortly thereafter, the wave nature of light was proposed and demonstrated by Thomas Young in 1801, based on which the phenomenon of wave interference was derived. Young was ably and experimentally supported by Augustine-Jean Fresnel, who contributed significantly in establishing the wave theory of light. Peter Ewart, in a paper published in 1813, gave evidence in support of the conservation of energy. Hans Christian Orsted observed in 1820 that a

magnetic force is developed by a conductor carrying an electrical current. At approximately the same time, it was discovered by Andre-Marie Ampère that two electric currents running parallel exert forces between themselves. In 1821, the analysis of the energy of a system expressed in terms of coordinates and momenta was initiated by William Hamilton, and his equations became known as the Hamilton's function. In the same year, an electricity-powered motor was built by Michael Faraday, while in 1826, Georg Ohm stated the law of electrical resistance, which expressed the relationship between current, resistance, and voltage in an electrical circuit. In the subsequent year, Brownian motion, as it came to be called, was discovered by the botanist Robert Brown. He hypothesized that the waterborne pollen grains that he observed under a microscope underwent constant movement due to collisions with the fast-moving molecules or atoms of which the water was composed. In 1831, the process of electromagnetic induction was discovered by Faraday, suggesting that magnetism can give rise to electric current or potential. The independent discoveries of Joseph Henry and Faraday on electromagnetic induction formed the basis for the inventions of the electric generator and the electric motor, respectively. Self-gravitating ellipsoids rotating uniformly were discovered in 1834 by Carl Jacobi. In the same year, John Russell found in Edinburgh's Union Canal a *soliton*, or solitary nondecaying water wave. He then went on to study the dependence of the wave velocities of solitary water waves on water depth and wave amplitude. The canonical equation of motion was proposed in 1835 by William Hamilton. In the same year, the Coriolis effect was deduced by Gaspard Coriolis after careful theoretical examination of the mechanical efficiency of the waterwheel. The effect, put simply, is that a rotating system experiences a force that acts perpendicular to the direction of motion and the axis of rotation. Hence, moving objects tend to be deflected to the right in the northern hemisphere and to the left in the southern hemisphere. A paper on conservation of energy was written in 1841 by Julius Robert but was rejected for publication. A few years later, in 1847, the principle of conservation of energy was independently proposed by Hermann von Helmholtz, who was credited with its formulation. In 1842, the Doppler effect, so called, which concerns the increase or decrease in the

frequency of sound, light, or other waves when source and observer move toward, or away from, each other, was stated by Christian Doppler. In 1851, Léon Foucault's pendulum showed the rotation of earth. The first half of the century saw significant advancements in continuum mechanics (e.g., for fluids), Navier-Stokes equations, and for solids the laws of elasticity. Of the many influences on the development of physics in the 19th century, the following are the most influential.

Heat

The self-taught physicist and instrument designer Guillaume Amontons is credited with the invention (1699) of the earliest air thermometer. Subsequently, with the thorough works of Victor Regnault (1841) and H. G. Magnus (1842) on the thermal compression and expansion of air, studies on thermometry came into prominence. The Carnot function (named for the engineer Nicolas Léonard Sadi Carnot), describing the relation between the amount of heat produced by a source of heat and the work that can be done by it, had been put to use in various ways by stalwarts like James Prescott Joule, Hermann von Helmholtz, and Benoit Paul Émile Clapeyron. The series of works on temperature and its measurements by William Thomson, Lord Kelvin, finally disposed of the subject. On the practical front, thermometry was significantly enriched through the efforts of Pierre-Louis Dulong and Alexis-Thérèse Petit (1818) through their work on the measurement of mercury expansion. This work was suitably supported by Regnault.

The memoir by Fourier of 1822 was responsible for the generation of the theoretical concept of heat conduction, which subsequently went on to illuminate much more of physics. This treatment has been successively dealt with by many of the mainstream thinkers of the time such as Kelvin, Siméon Denis Poisson, and others. It was Kelvin whose efforts resulted in the origin of the concepts of ingenious methods of sinks and sources. Following Fourier, greater attention was paid to earth's heat conduction. On the experimental side, in 1853, Gustav Wiedemann and Rudolph Franz could determine the relative heat conduction for the metals. They were able to show the existence of parallel gradation for the conductivity of electricity and heat. The works of Anders Jonas Angstrom, Franz

Ernst Neumann, and James David Forbes were instrumental in establishing heat conduction measuring absolute methods. Heinrich Friedrich Weber in 1880 devised a lamellar method that can be applied to fluids.

In 1763, the research work of Joseph Black was important in the completion of the work on practical calorimetry. The concepts of chemical, latent, and specific heats were experimentally enriched by the valuable work of Regnault. The approximation of atomic heat constancy for the elements was discovered in 1819 by Dulong and Petit. In 1875, Weber went on to interpret the apparently exceptional cases for boron, carbon, and silicon. This was followed by Regnault in 1840 for sulfur. In 1831, the law was applied to compound bodies by Neumann. Joule in 1844 was able to show that the component specific heats in many cases can be treated additively with the specific heat. Particular mention should be made of Robert Bunsen's ice calorimeter of 1870.

Radiation

With the thermopile reaching the hands of the Italian physicist Macedonio Melloni, the concepts of thermocrosis, diathermacy, and heat radiation gained prominence. The questions on heat and light identity were forever put aside by these researches. In 1878, the invention of Samuel Pierpont Langley's bolometer put the subject to a greater height with regard to precision. The laborious attempts of E. Becquerel and H. Becquerel and John Herschel on the survey of heat spectra ultimately led to the formulation of Langley's charts. The whole subject was perhaps pervaded in 1860 by Gustav Kirchhoff with his Kirchhoff's laws. These laws in their earliest probable forms were anticipated by the Scottish physicist Balfour Stewart as early as in 1857. Josef Stefan in 1879 tentatively formulated black body radiation in relation to temperature. In 1884, Boltzmann went on to further strengthen this. Among other inventions worth mentioning in this regard are Alexander Graham Bell's photophone and William Crooke's radiometer. The radiometer had been put to use by many important figures like Ernest Fox Nicholas and Gordon Ferrie Hull, James Clerk Maxwell, and Pyotr Nilolayevitch Lebedev in his actual measurement of light pressure. In 1838,

Claude Pouillet came up with the first ever estimation of solar radiation constant at earth. Angstrom in 1886 and Langley in 1884 went on to devise other pyrheliometric methods.

Thermodynamics

The vis viva of 1686 constituting the old Leibnizian principle was broadened and fruitfully interpreted through thermodynamics. To begin with, the experiments of Humphrey Davy (1799), Benjamin Thompson, and Count Rumford (1798) set the tone for the domain of thermodynamics. It almost leaped with the computation of mechanical equivalent of heat by Julius Robert von Mayer, and an epoch was marked when Joule came up with the series of judiciously varied and precise measurements of heat (in 1843, 1845). Helmholtz proposed and spread his theory of conservation of energy to encompass the whole of physics following these discoveries. Prior to this, the early part of the century saw Carnot inventing the reversible thermodynamic cycle. The theory of equilibrium for two bodies in mutual action, or the reversible thermodynamic cycle, made huge contributions to modern physics. The efficiency of engines derived from the point of temperature criterion was deduced by Carnot. The geometrical method of representation as proposed by Clapeyron is used today universally. However, Josiah Willard Gibbs with his suggested new coordinates in 1873 allowed it to be used more flexibly. But more work needed to be done in defining the value of transformation so as to bring in harmony between the first law of thermodynamics and the ideas of Carnot. This was just what Rudolf Clausius did in 1850, supported by Kelvin and William J. Macquorn Rankine in the following year. Rankine's broader energetics treatment has been the basis of many subsequent discussions. The investigation on the thermodynamic solution of solids and gases was initiated as early as 1858 by Kirchhoff, who introduced an original method of treatment. Most of the objections to the second law of thermodynamics were dispensed with by Clausius, who was successful in defending even the objections associated with radiation. The law that came nearest in logically justifying the second law of thermodynamics is the elementary statistical mechanics principle of Gibbs, which he proposed in 1902. The first and second laws of

thermodynamics have ubiquitous applications. Some of the interesting instances include the concept and properties of an ideal gas, Kelvin and Joule's experiment to show the physical departure from ideality. The difficulties encountered by Louis Paul Cailletet and Raoul Pictet in liquefying incompressible gases were not faced by James Dewar and Morris William Travers or Karol Olszewski for hydrogen coercions. Further examples include broader treatment of evaporation and fusion, from the calculation of under pressure melting point of ice by James Thomson in 1849, treatment of sublimation by Kirchhoff in 1858, and Kelvin's equation (1883) based vast chapter on thermo-elastics. Besides, there were discoveries of liquid and gaseous state continuity in 1869 by Thomas Andrews; Émile Amagat's experimental prowess in furnishing deeper insights into the laws of physical gases and Johannes Diderik Van der Waals's great work of approximation in 1873 that almost closely predicted the observed facts. The further extension and development of thermodynamics was yet to take place that mainly concerned the chemical systems. Many thinkers at the same time recognized the thermodynamic potential and its analytical power. Some of the important contributions were made by Gibbs, who in 1876 proposed for both adiabatic and isothermal potential, the characteristic functions of Helmholtz in 1882. Gibbs made the most of the new research possibility originating with complete contrast to virtual displacement of *mechanique analytique* and the introduction of the modification of virtual thermodynamics. Without the strong mathematical forces, Jacobus Henricus van 't Hoff came up with his simple and marvelous application of the second law of thermodynamics and propounded a new solution theory based on interpretations of experiments of François-Marie Raoult and Willhelm Pfeffer. This became the basis for experimental investigation of chemical physics. Gibbs's highly generalized chemical statistics brought early success when applied to the phenomenon of dissociation proposed by Henri Sainte-Claire Deville in 1857, and subsequently, it was followed by Gibbs, Pierre Duhem, and Max Planck deducing in succession appropriate equations. The law of mass action (1879), based on research by Cato Guldberg and Peter Waage, received a theoretical basis in case of dilute solutions through the works of Planck.

Elasticity

The theory of elasticity that has made longstanding contributions to physics and has become integral throughout the whole discipline can be traced to a 19th-century creation. In 1807, Thomas Young had recognized shearing force and brought about the useful concept of modulus. Earlier experimental indications had come from Charles-Augustin de Coulomb (1776) in his attempted analysis of torsion and flexure; Leonhard Euler and Jacob Bernoulli's (1742) treatment of vibrating bars, Bernoulli's proposal of the elastic curve along with the works of Edme Mariotte (1680), Robert Hooke (1660) and Galileo Galilei (1638). Claude-Louis Navier's equations of elastic potential energy and elastic displacement enhancing the molecular hypothesis commenced around 1821 that on the broader lines marked the beginning of establishment of the elasticity theory. However, Augustin-Louis Cauchy's (1827) brilliant generalizations soon discredited this startling advancement. Among the valuable contributions he made to the field are strain quadric and stress quadric, six component strains and six component stresses. He also happened to reduce the components to three principal strains and three principal stresses. He explained the ellipsoids and some of the most indispensable concepts of the elastic physics of modern day. Cauchy came to his conclusions and equations aided by molecular hypothesis as well as oblique stress analysis across an interface. These methods gave rise to 15 elasticity constants in the most general case. In the similar time, simultaneous independent contributions came from Poisson (1829) and Clapeyron (1828). In 1837, George Green introduced another fundamental and independent method in elastics. He tried to explain the potential energy of the closed system based on the virtual displacement principle of Joseph-Louis Lagrange. This method was further enhanced through the contributions of Neumann (1885), Kirchhoff (1876), and Kelvin (1856) leading to the formulations for the aeolotropic medium, 21 constants which could be reduced to two for the simplest case. In 1828, Poisson first deduced isotropic medium wave motion. He happened to show the occurrence of different velocity constituted transverse and longitudinal waves. The wave motion problems treated by Green in 1842 for the

aeolotropic media came under significant and strong criticisms from Christoffel in 1877 and Blanchet (1840–1842). The case of radial vibrations for a sphere was treated by Poisson in 1828. Alfred Clebsch in his work of 1862, *Vorlesungen*, generalized in solids, the free vibrations theory. The whole subject was perhaps overhauled with longstanding and continual efforts of de St. Venant (1839–1855), providing the major thrust to elasticity. His works were supported by Clebsch in 1864 and by Navier and François Moigno in 1863. The fundamental significance of shear was adequately asserted by him. In 1855 and 1856, the flexure and torsion of prisms in their complete form was provided through the profound and intricate researches of Adhemar Jean Claude Barré de St. Venant. In either case, the stressed-solids right sections were shown to be curved wherein the curvature could be specified succinctly. The first case superseded the inadequate torsion formula of Coulomb, and the latter case reduced flexural stress to a coupled and transverse force. Among the other notable contributions to the field of elasticity made in the 19th century came in 1852 that concerned the work of Gabriel Lamé for the application and invention of curvilinear coordinates. In 1872, Enrico Betti proposed the reciprocal theorem. Valentino Cerruti in 1882 applied this theorem to plane boundary solids. The same problem was treated by other methods by the likes of Joseph Valentin Boussinesq, Lamé, and Clapeyron. Further support came through notable contributions from Lamé's study on strained spheres in 1854, the treatment of flexed plates in 1850 by Kirchhoff, treatment of oscillating systems with finite degrees of freedom by Rayleigh in 1873, and Duhamel's (1838), Kelvin's and Neumann's equations on thermoelasticity. Kelvin moreover made some important contributions with his works in 1863 on the elastic spheroids and their geophysical applications as well as with his study of torsion of prisms in 1878.

Dynamics

Joseph-Louis Lagrange's unrivaled feat of reasoning of the 18th century, his *Mecanique Analytique* of 1788, was inherited by the following century's investigators of pure dynamics. Until 1813, the presence of the master played an important part

in watching over the legitimacy of its applications and pointing out its resources. Each new advancement made had gone on to vindicate the preeminent safety and power of its methods. There were successes with the deduction of electromagnetic field kinetics by Maxwell in 1864 and its adaptation to the chemical systems equilibrium by Gibbs. The method of *liasons* faced natural reactions. The most outspoken proponents came under the protection of Laplace, for example, Poisson's *Mecanique Physique* of 1828, an extension of dynamics of Roger Boscovich which permeated the works of such figures as Frensel, Navier, Cauchy, Ampère, Boussinesq, de St. Venant, and others. Cauchy had tried to reconcile molecular methods based on Lagrangian abstractions. Poisson's method proved to be untrustworthy, however, notwithstanding its superiority in the instances of classic elastics and capillarity. The influence of the medium came to the fore with the rise of modern electricity, and Poisson's method was rudely shaken.

Thomson and Peter Tait are credited with a complete revision of dynamics during their endeavor to obtain uniformity and clarity of design by going back to Newton and logically reconstructing the whole field. The principle of conservation of energy was systematically used for the first time in this major work. Going with the contemporary trend, Hertz in 1894 made a strong effort to exclude altogether potential energy and force from the purview of dynamics. He went on to postulate a universe of concealed motions as in Helmholtz's treatment of cyclic systems theory and the idea of a dynamic gyrostatic ether proposed by Kelvin in 1890. In fact, Boltzmann and Helmholtz introduced ordered molecular motions, and concealed systems proved to be highly potent in the justification of Lagrange's dynamics, particularly when applied to actual natural motions.

During the 19th century, first-rank-specific contributions with mathematical difficulties or in the applications of the subjects were very few. The principle of least constraint proposed in 1829 by Carl Friedrich Gauss, although superior to the displacement method in its applications, was seldom used. But it was the development and adaptation of Gottfried Wilhelm Leibniz's (1646–1716) systematic formulation of first principles to the law of conservation of energy that impelled the

progress of the domain of dynamics in the 19th century. Although the writings of Lagrange in 1777 included an explanation of potential, and Laplace in 1782 had already provided the differential equation on which the principle was based, the latter was completed by Poisson only decades later, in 1827. Major contributions came from Gauss with regard to force flux and surface integrals when he proposed his influential theorems in 1813 and subsequently in 1839. In 1854, George Stokes contributed important work on the relation of surface and line integrals. The most prominent of the authors on volume surface and line potential were Michael Chasles, Robert Murphy, Peter G. L. Dirichlet, C. Neumann, and Helmholtz. Gradually, the dynamists of subsequent generations have shown more inclination toward vector calculus and geometry. A significant contribution in this regard was made by August F. Möbius. In addition, William Rowan Hamilton's *quaternions* of 1853 and Hermann Grassmann's *Ausdehnungslehre* ("On Expansion") of 1844 played vital roles. Another work of major significance was Sir Robert Ball's 1876 study of the theory and immediate dynamic applications of screws. Other work with profound experimental aspects include Hippolyte Fizeau's gyrostat and Léon Foucault's pendulum, both of 1851, and the quartz-fiber torsion balance of Charles Vernon Boys (1887) that could measure, with greater precision than any other apparatus, the earth's mean density and Newton's gravitational constant. Some other important work of this period includes George Biddell Airy's pendulum of 1856, Philipp von Jolly's balance methods of 1881, and the contributions of Arthur König and Franz Richarz, published in 1885. The century also saw the beginning of Thomas C. Mendenhall's extensive U.S. transcontinental surveys in 1895.

Hydrodynamics

The extensive labors of Lagrange, Alexis Clairaut, and Colin Maclaurin between 1742 and 1788 made it possible to understand the equilibrium of liquids even with regard to rotating fluids before the beginning of the 19th century; the domain of hydrodynamics thus bequeathed a rich legacy of development. The triaxial ellipsoid of revolution was generalized by Moritz von Jacobi in 1834.

George Howard Darwin (1887) and Henri Poincaré (1885) extended it to two rotating attracting masses. Particular mention should be made about the astonishing revelations in the works of Poincaré. The subject was divided into classes of motion (flow in jet waves, streams, and tubes) possessing velocity potential by Helmholtz in his famous paper of 1858. Thomas Craig and Henry A. Rowland in 1881 carried out the classification for the higher orders of motion. The century saw significant progress in the treatment of waves for cases with velocity potential, dynamics of solids in frictionless liquids, and the treatment of discontinuous fluid motion. Progressive wave motion in shallow vessels was dealt with by Green, Scott Russell, and Philip Kelland, while those for deep water progressive waves were studied by Rankine and František Josef Gerstner. Most of the problems faced in the advancements of hydrodynamics were countered by the analytical works of Stokes. The case of ripples, introduced by Kelvin in 1871, was treated by John William Strutt, Lord Rayleigh, in 1883. The efforts of Dirichlet, Stokes, and Poisson solved the problem of spheres, while the problem of ellipsoids was touched upon by Kirchhoff, Clebsch, and Green. Rankine was successful in his treatment of ellipsoidal and cylindrical translatory motions in a way similar to that of the resistance offered by ships. Bjerknes's results with the hydrodynamic force around spheres including rotational, translatory, oscillating, or pulsating have tremendously enriched the field of hydrodynamics. Other influential work in this domain include Greenhill's treatment of ballistics and Josef Ressel's design of the first successful screw ship propeller. The question on the stability of the motion of fluids has been highlighted considerably by the numerous contributions of Kelvin. The study of vortex motion was approached gradually by the likes of George Stokes, Cauchy, and Lagrange. But it was Helmholtz's integration of the differential equations that made for clear interpretations of the whole matter.

Optics

Prior to the 19th century, valuable contributions by Lagrange, Étienne-Louis Malus, Newton, and Christiaan Huygens led to much progress in geometric optics. Gauss, in 1841, put forward a

paper that remodeled the topic and was generalized subsequently by Ernst Abbe and Möbius. In 1835, George Biddell Airy first pointed out the role of diffraction on geometric optics, which was further developed by Rayleigh and Abbe. One of their major accomplishments is the theory of the rainbow. In 1879, Abbe introduced the apochromatic lens, which offers greater correction of chromatic and spherical aberration than the more common achromatic lenses. Jesse Ramsden and Huyghens's eyepiece, which reduces blurring from aberrations, still served the microscope well; however, successive stages of improvement took place in the objective lens, beginning with the invention of the achromatic objective lens by Joseph Jackson Lister in 1830. A series of developments in geometric optics took place, like J. W. Stephenson's introduction of the principle of immersion (1878), Abbe's homogenous immersion (1879), and the apochromatic objective of Abbe-Zeiss (1886). These microscopes gained significantly from the experiments conducted in the glass factory in Jena, Germany. The photographic objective's guiding principle was introduced by Steinheil between 1865 and 1866. The difficult task of constructing fine refractive telescope lenses for major astronomical observatories was carried out by the firm of Alvan Clark & Sons in Cambridge, MA, founded in 1846; many of these telescopes are still in use.

Newton's work nearly culminated in the achievement of spectrum purity early in the 18th century. Nearly a century later, in 1802, it was William Hyde Wollaston who introduced the concept of slits that enabled the visualization of solar spectrum dark lines. Their origin was carefully mapped by Joseph von Fraunhofer, between 1814 and 1823. The first ever notion about the occurrence of spectral absorption was stressed by David Brewster, who believed the phenomenon to be atmospheric. In 1860, he also published a map of 3,000 spectral absorption lines. The D group of Fraunhofer coincided with the sodium lines; this fact was pointed out by Lèon Foucault in 1849, and thereby, sodium vapor's reversing effect was unraveled. The works of Stewart in 1860 and Angstrom in 1855 were instrumental in describing the parallelism of absorption and emission. However, Kirchhoff and Bunsen in between 1860 and 1861 ultimately described the differences between

the dark lines or fixed bright line spectra and the continuous spectra on which basis spectral analyses are made. Later on, improved observations on the wavelength and distribution of solar spectra were made by Robert Thalén, Angstrom, and William Huggins. A definite advancement was noticed with each of the discoveries of Rowland, such as his solar-spectrum maps and tables of 1887 and 1889, and his summary in 1891 of the sun's elemental composition. Alkaline and other elemental structural spectra were successfully analyzed by Kayser and Runge. In 1874, Rayleigh introduced the modernized version of grating theory, which got extended to the concept of concave grating, particularly through the works of Rowland. Langley's 1883 investigation of the infrared spectrum paralleled Rowland's works on visible spectrum. The extension of the detailed accuracy of the spectrum was brought about by the works of Langley, which superseded the investigations of earlier workers like Fizeau and Foucault. The accuracy of the spectrum got improved to nearly eight times of its visible length. The radiations of hot and incandescent bodies, the lunar and solar spectrum were all precisely specified. Our knowledge of solar activity was highly enriched through the valuable experimental suggestions of Lockyer's 1868 spectroheliograph and Jansen's works and the remarkable success achieved though it by the likes of George Ellery Hale.

Playing a major role in the acceptance of the wave theory of light was the work on interference as framed by Thomas Young in 1802, which advanced the earlier experimental work on colors of thin plates by Newton, Hooke, and Boyle. The colors of mixed plates also were investigated by Young, who was aware that reflection from a denser medium results in the loss of nearly half a wavelength. In 1819, Newton's colors were independently explained by Augustin-Jean Fresnel, a civil engineer, in terms of wave interference. Through the course of this work, he devised his triple mirror, biprism, and double mirror. Albert A. Michelson went on to make crucial advancements to the scope of this device, which became a fundamental tool for research in this field. At the U.S. Naval Observatory in 1877 and 1879, Michelson estimated the speed of light in air and in a vacuum with great accuracy. An entirely new interference phenomenon was introduced by Otto

Weiner in 1890 when he discovered the stationary light wave. In the 1890s, Gabriel Lippmann used a two-step method successfully to reproduce colors in photography, for which he would receive the 1908 Nobel Prize in Physics. Understanding of the broader aspects of interference, such as multiple reflections, was successfully advanced by the works of Airy, Fresnel, and Poisson.

Syed Feroj Ahmed

See also Physics, 20th century; Physics, Antiquity; Physics, Contemporary; Physics, Enlightenment; Physics, Middle Ages; Physics, Renaissance; Physics, Thermodynamics

Further Readings

- Airy, G. B. (1835). On the diffraction of an object-glass with circular aperture. *Transactions of the Cambridge Philosophical Society*, 5.
- Batchelor, G. K. (1967). *An introduction to fluid dynamics*. Cambridge University Press.
- Boresi, A. P., Schmidt, R. J., & Sidebottom, O. M. (1993). *Advanced mechanics of materials*. Wiley.
- Cahan, D. (Ed.). (1994). *Hermann von Helmholtz and the foundations of nineteenth-century science*. University of California.
- Dingle, H. (1963). A hundred years of spectroscopy. *British Journal for the History of Science*, 1, 199–216.
- Fresnel, A. (1818). Memoir on the diffraction of light. In *The wave theory of light: Memoirs by Huygens, Young and Fresnel*. American Book Company.
- Gibbs, J. W. (1873). A method of geometrical representation of the thermodynamic properties of substances by means of surfaces. *Transactions of the Connecticut Academy of Arts and Sciences*, 2.
- Griffiths, D. J. (1999). *Introduction to electrodynamics* (3rd ed.). Prentice Hall.
- Kirchhoff, G. (1860). Über das Verhältniss zwischen dem Emissionsvermögen und dem Absorptionsvermögen der Körper für Wärme und Licht. *Annalen der Physik und Chemie*, 109(2). Translated by Guthrie, F., & Kirchhoff, G. (1860). On the relation between the radiating and absorbing powers of different bodies for light and heat. *Philosophical Magazine*, 20(4).
- Maxwell, J. C. (1965). A dynamical theory of the electromagnetic field. *Philosophical Transactions of the Royal Society of London*, 155.
- Moyer, D. F. (1977). Energy, dynamics, hidden machinery: Rankine, Thomson and Tait, Maxwell. *Studies in History and Philosophy of Science*, 8(3).

- Petit, A. T., & Dulong, P. L. (1819). Recherches sur quelques points importants de la Théorie de la Chaleur. *Annales de Chimie et de Physique*, 10. *The Philosophical Magazine: A Journal of Theoretical Experimental and Applied Physics*, 33, 34(1867).
- Rozenfeld, B. A. (1988). *The history of non-Euclidean geometry: Evolution of the concept of a geometric space*. Springer.
- Volta, A. (1800). On the electricity excited by the mere contact of conducting substances of different kinds. *Philosophical Transactions of the Royal Society of London*, 90.
- Young, T. (1845). *Course of lectures on natural philosophy and the mechanical arts*. Taylor and Walton.

PHYSICS, QUANTUM THEORY

Quantum theory is the branch of physics dealing with the behavior of matter at the atomic and subatomic level. It developed during the period 1900–1930, through problems arising at the intersection of thermodynamics, electromagnetism, and the discovery of atoms, electrons, and radioactivity. It has been an astonishingly wide-ranging and successful scientific theory. It has been tested in experiments on energy scales comparable to the early stages of the universe; it has provided explanations of stellar and cosmological phenomena; it has underpinned advances in materials science and chemistry ranging from the development of semiconductors in computing to biological processes at the cellular level.

Its key characteristic is that properties of atomic and subatomic objects can only occur in, or change by, discrete quantities, the smallest of which is a *quantum*. The consequences of this effect have been extraordinarily wide ranging: Properties can only be measured imprecisely; many combinations of properties cannot be simultaneously measured; systems cannot be prepared with well-defined values for many properties; the results of measurements need to incorporate randomness in their description; and the properties inferred from the outcomes of measurements appear to become properties of the manner of the measurement process as well as properties of the object itself.

Despite its empirical success, quantum theory remains one of the most puzzling of scientific theories and has had a significant impact upon the philosophy of science. The question of how one interprets quantum theory remains debated. It appears to suggest the world is very different from how we might have imagined it: ideas such as nonlocal connections between remote locations; branching multiverses; and whether physics should be understood as attempting to give a description of reality at all, have all been advanced in attempts to understand how it can be that quantum theory correctly describes the way the world works.

This entry traces the development of quantum theory in the early 20th century, continues with a brief characterization of the underlying mathematics, and describes the dual nature of quantum objects as both wave and particle. Next discussed are the uncertainty of assigning values to multiple properties of quantum objects, the ambiguity of measurement in quantum theory, and the nature of nonlocal connections. The entry concludes with an explanation of empirical and realist interpretations of quantum theory.

Historical Development

Quantum theory grew out of developments in late 19th-century physics. The *classical* physics of Newtonian mechanics described particles moving in space under the influence of gravitational, electric, and magnetic fields. A field is a property, defined at every point in space, that describes the force that will act on a body at that location. The unification of the electric and magnetic fields by James Clerk Maxwell (1831–1879), in such a way as to show that light could be identified as a wave in the electromagnetic field, had been a triumph of the Newtonian paradigm.

However, the attempt to apply thermal physics to the electromagnetic field seemed to lead to a problem. If a system was able to interact with the electromagnetic field, all of its heat would be absorbed into increasingly high frequency waves of the electromagnetic field. In 1900, Max Planck suggested the problem could be solved if the energy in the electromagnetic field could only be exchanged with matter in packets of a fixed size: a minimum *quantum* of energy, determined by a new universal constant. This quantity increased with frequency,

so the electromagnetic field would be blocked from absorbing energy at extremely high frequencies.

Albert Einstein, in 1905, showed that this light quantum proposal could also account for a phenomenon known as the *photoelectric effect*. Light shone on a metal can cause electrons to be emitted, as the electrons absorb the energy of the light. This effect ceases if the light is below a threshold frequency, regardless of the intensity of the light. Einstein explained this, in terms of a minimum quantum of energy needed to eject the electron from the metal. Unlike Planck, who only suggested that energy must be *exchanged* in discrete quantities, Einstein suggested that the electromagnetic field, including light, was *composed* of individual particles, or light quanta.

At the same time, the discovery of radioactivity and electron-scattering experiments started to provide evidence for subatomic structure: Atoms seemed to be concentrated heavy nuclei of positive charge surrounded by light, negatively charged electrons. According to the theory of electromagnetism, such structures would be highly unstable and would collapse, with the electron spiraling into the nucleus. In 1913, Niels Bohr showed that if the quantum hypothesis was applied to the angular momentum of electrons orbiting the nucleus, then the electrons could only inhabit fixed orbits, preventing the instability. Not only would this solve the problem but jumps between the stable orbits would absorb and emit light quanta with characteristic frequencies (Balmer lines) that had already been observed in hydrogen.

The synthesis of the two problems came in 1924. Planck and Einstein's hypothesis had suggested that light, previously considered a wavelike entity, showed particle-like properties. Louis de Broglie showed that if matter, previously considered a particle-like entity, had wavelike properties, then Bohr's results could be derived.

Two rival mathematical versions of quantum theory developed to describe this: matrix mechanics was developed by Max Born, Pascual Jordan, and Werner Heisenberg, which tended to emphasize discontinuous quantum jumps between stable states of quantum systems; and wave mechanics, developed by Erwin Schrödinger, which originally attempted to describe all quantum processes as continuous wavelike effects. The reinterpretation of Schrödinger's wave equation in terms of

probabilities, by Born, paved the way for Paul Dirac, in 1930, and John von Neumann, in 1932, to show how the two approaches could be unified as examples of a more general structure, called *Hilbert space*, after the influential mathematician David Hilbert. Mathematically, the structure of quantum theory was essentially finalized. The problem of understanding the picture it gave of the world remained.

The Mathematical Structure of Quantum Theory

In classical Newtonian physics, the mathematical structures used are often understood as modeling the way the world actually is. The state of a system is just the list of all the physical properties possessed by the system, and the dynamics describe how those properties change over time.

When describing a system using quantum theory, the quantization process starts with a classical model for a system, then replaces the mathematics with equivalent structures in a Hilbert space. A property is represented by an operator, and a state becomes a vector in the Hilbert space. However, in the mathematical structure to which this leads, the state no longer lists a specific value for each property. Instead, it only gives the probability that the property may have a given value. The dynamics, describing how the state changes in time, gives how those probabilities change in time.

Nevertheless, whenever a property of a quantum object is observed, the observed property will have a well-defined value. Unlike in classical physics, therefore, the quantum state does not seem to be a mathematical representation of the way the world is. Instead it only gives the probabilities of what can be observed, if an observation is made. For this reason, the properties of quantum theory are typically referred to as *observables*, emphasizing the fact that they represent the outcomes of observations, rather than properties of the object itself.

The mathematical structures of quantum observables are very different from classical properties: There are strong limitations on the extent to which quantum properties can be observed, and on which quantum properties can be observed simultaneously. Quantum objects at remote locations seem to show instantaneous influences on each other's properties. The question remains as to whether

quantum objects can possess properties, and of what kind, when they are not being observed.

In prequantum physics, the mathematical description of physical theories seemed to naturally represent physical processes. They came equipped with a picture of the way the world could be, that seemed to explain why the mathematics was successful at making predictions. The mathematics of quantum theory does not appear to provide such an explanation, and providing this is the task of interpreting quantum theory.

Wave–Particle Duality

Quantum objects, whether as electromagnetic fields or as electrons, show both wavelike and particle-like properties, depending upon how they are observed. Experiments using beam splitters can be used to show these properties.

A beam, of particles or waves, passes through a splitter, which deflects half the beam into a second path. The split beams are then reflected and recombined into a single beam. Detectors can be placed into the paths of the beams, to measure the intensity of the beam in any given location.

When the beam is made of particles, and the intensity is reduced to a very low level, the detectors will start to register individual particles. They can track which path the individual particles take, and the detectors will register a particle taking one path or the other, but not both. The intensity of the recombined beam is always equal to the sum of the intensities from the two split beams.

When the beam is a wave, like a water wave being driven down a narrow channel, detectors placed in the two paths will continue to register waves occurring in both channels, regardless of how much the intensity of the wave is reduced. When the beams are recombined, waves can exhibit interference: The intensity of the recombined beam does not necessarily add up to the sum of the intensities of the two split beams. Where the intensity increases, it is called *constructive interference*, and when it decreases it is *destructive interference*. The nature of the interference can depend upon interactions with the waves in both channels.

When a quantum beam is split into two different paths and recombined, constructive and destructive interference can occur in the intensity of the recombined beam. The scale of this interference depends

upon interactions in both paths. This is a signature of waves traveling down both paths. However, when detectors are placed in both paths, at low intensities individual quantum objects are detected in one path or the other. This is a signature of a particle, traveling down only one path.

When the particle-like direction taken by a quantum object is detected, the wavelike interference effects disappear from the recombined beams. Paradoxically, if the detectors are removed from the split beams but the intensity is still kept at the very low level, where it seems individual quantum particles must be following one path or another, wavelike interference effects reappear in the intensity of the recombined beam. It seems as if the quantum object behaves as a particle when observing for particles and as a wave when observing for waves.

Quantum Uncertainty

In classical physics, objects were assumed to have well-defined values for all their fundamental properties, such as mass, electric charge, and so forth. In principle, these properties could all have been simultaneously measured and observed to have well-defined values.

In general, a quantum state only describes probabilities of observing values on measurement. In some cases, though, a quantum state gives a probability of one to a particular result, and then the property does have a well-defined value. In particular, if a property has been observed to have a particular value, the quantum state must change, to reflect the observed value.

However, some combinations of quantum observables cannot be simultaneously assigned well-defined values. The paradigmatic example is the position and momentum of an object. There are no quantum states that can give a well-defined value to both properties. In general, these complementary observables will have a minimum joint uncertainty, so that when one is very likely to be found within a narrow range of values, the other will have a very large spread of possible values. When an observable is measured, and the quantum state changes to reflect the observed value, the new state must introduce a large uncertainty in the value of any complementary observable.

The naive view of this joint uncertainty, much discussed in the development of quantum theory, was that the measurement process, needing to exchange a minimum quantum when interacting with the system, introduces a random disturbance to the complementary properties.

However, the joint uncertainty can be seen without introducing such measurement disturbance. Suppose a collection of quantum objects are prepared in the same quantum state. Half are randomly selected for observation of one property, and half have the complementary property observed. None of the objects is subjected to both measurements. Nevertheless, the spread in values over the two sets still shows the joint uncertainty: If the first group reveals only a narrow range of values, the second group must show a wide spread in the complementary property, and vice versa. The joint uncertainty is already present in the quantum state, before measurement takes place.

The Measurement Problem

The quantum state doesn't attribute definite values to properties. Quantum objects are only described by the probabilities that values will be observed when measured. However, quantum theory leaves ambiguous what constitutes a measurement.

A measurement process involves a quantum object interacting with a much larger device, which records the outcome. Typically, the measuring device will be a much larger system, amplifying the outcome up to a level that is directly observable, such as numbers printed on paper or a screen.

The measuring device itself, however, is built up out of atoms and their interactions, so presumably quantum theory should also apply to the measuring device itself. This leads to a problem: when the measuring device is also described by quantum formalism, it does not attribute definite values to the measuring device either, including not giving a definite value for the outcome of the experiment itself. Most problematically, the quantum state seems to allow the possibility of constructive and destructive interference effects between the different possible outcomes of the measurement. This is characteristic of wavelike properties and seems incompatible with the idea that the measuring device really has recorded one outcome or another. The quantum formalism only gives a probability

that the measuring device will show a given outcome when the device itself is measured. Attempting to measure this device with another measuring device only shifts the problem on, as trying to describe the new device using the quantum formalism raises exactly the same issues.

Real measuring devices show a statistical mixture of definite outcomes, which cannot, even in principle, exhibit interference. The difference between the quantum state—which describes no particular outcome and maintains the possibility of interference—and the statistical mixture of outcomes has been described as the “collapse of the wavefunction.” The single quantum state collapses onto a probability distribution over quantum states, each one representing a different possible outcome. The problem for quantum theory is that no process within the quantum formalism can rigorously describe such a collapse.

Quantum Nonlocality

Recall that in the beam splitter the atom appears to go down one path, while still being affected by interactions in the other path. Although this seems paradoxical, there is no detectable evidence that interactions in one path are having an immediate influence in the other path: the observable results only appear when the beams are recombined. If there were instantaneous influences, then this would be a significant problem for the theory of relativity, which forbids such influences.

When two quantum objects are allowed to interact and separate, however, the situation becomes more problematical. In 1935, Albert Einstein, Boris Podolsky, and Nathan Rosen argued that strongly correlated quantum objects must either use instantaneous influences, or else the objects must actually have definite values for their properties even though the quantum state does not describe them.

A strongly correlated (known as *entangled*) quantum state of two objects can ascribe a definite value to the joint properties of the system, such as the sum of their momentum, without ascribing a definite value to the individual properties of the subsystems, such as the individual momentum. If the objects are widely separated, and the momentum of one object is measured, it will acquire a definite value. However, the second object also

must acquire a definite value of momentum, as the sum of the two remains fixed.

Either the value of the second momentum was *already* definite prior to the measurement occurring, or its value was *made* definite by that remote measurement. The first possibility means that the second object must have had a definite momentum, even though the quantum state did not describe it. The second option implies the measurement had an instantaneous influence on a remote location.

Problematically for Einstein and colleagues' argument, in 1964 John Bell was able to show that the correlations between remote quantum objects could not be explained without instantaneous influences. Einstein had sought an explanation for the correlations between the remote objects entirely in terms of the existence of properties prior to the remote measurements being performed. Bell showed that this implied a limitation on how well the two objects could show correlations for combinations of different measurements. Entangled quantum states are able to exceed this limitation. The conclusion would appear to be that measurements can create or influence properties at remote locations, regardless of the conflict with relativity theory.

Empiricist Interpretations of Quantum Theory

Empiricist approaches to the problem of interpreting quantum theory focus upon the mathematics of quantum theory as a tool for making predictions and suggest that is all a good scientific theory should aspire to. Further questions about what might be happening when a system is not being observed are regarded as pseudoproblems, without empirical content.

The Copenhagen interpretation, as it is known, was the first such approach, which was actually a number of distinct ideas associated with Bohr and Heisenberg. It was strongly influenced by positivism, and by ideas that it was meaningless to attribute values to quantities that are not directly measured or observed. It came to dominate the interpretation of quantum theory for several generations.

A key element was Bohr's analysis of the process of measuring objects at the atomic level. Because atomic objects cannot be directly observed,

this always requires some amplification process, to bring the result up to a level where a classical description becomes valid. The interaction between classically describable experimental devices and an atomic object necessarily involves an exchange of discrete quanta. When the interacting systems are both large, the scale of this interaction can be treated as negligible with respect to the properties being observed. As the system of interest became smaller, the minimum quantum of interaction between the systems means that it was no longer possible to treat the interaction of the measuring device with the system as negligible. The phenomenon that emerges in the measuring device is a property of the whole arrangement of system and measuring device and cannot be attributed solely to the quantum object; Bohr argued that attempts to do so lead to paradoxical properties, as exemplified by wave-particle duality. Bohr's position emphasized the holistic aspect of the process, so that particle-like phenomena are properties of experiments for particle-like properties, and wave-like phenomena are properties of experiments for wavelike properties, associated with two different types of experimental apparatus. The paradox, to Bohr, only seems to arise if one tries to attribute the wavelike or particle-like properties to the quantum object itself, independently of the means used to measure it.

This position does not necessarily deny the existence of quantum objects or that they may possess properties, but it does not regard the properties predicted by quantum theory as pertaining to the quantum objects alone. Quantum theory can only discuss the properties of experimental arrangements, without further analysis into properties of subsystems.

This position dissolves the measurement problem by suggesting it was a mistake to apply the quantum formalism to the measuring device. The existence of the directly observable, experimental device has to be taken for granted. The quantum formalism is not a description of the properties of quantum objects but instead is a tool for predicting the behavior of the experimental apparatus when its interactions are sensitive to the individual quanta exchanged between systems.

Quantum Bayesianism is an example of a modern empiricist interpretation. This holds, broadly, that quantum theory is a form of inferential

reasoning, generalizing classical Bayesian probability theory. As a Bayesian theory of probability, it takes the view that the probabilistic statements quantum theory makes are strictly the degrees of beliefs, exemplified by gambling commitments, of rational agents, as to the outcomes of their observations.

The empiricist point of view is compatible with a number of different attitudes to the existence of quantum objects and their properties: They may exist, but it is impossible to describe them, or meaningless to talk about them or analyze their properties; or it is incoherent to hold beliefs about their properties that go beyond beliefs about the empirical predictions one can make about the results of interactions with quantum objects. The most radical claims, that there is no world independent of our interactions with it, moves away from empiricism and into skepticism or solipsism.

Realist Interpretations of Quantum Theory

Realist approaches to quantum theory accept the same empirical content of quantum theory as the empiricist but believe it is still fruitful to ask how the world could be so that quantum theory holds. Three broad perspectives have developed within this.

The first point of view attempts to understand the mathematical formalism of quantum theory as a literally true description of the way the world is, without addition or modification. The most immediate challenge to this is that of accounting for the measurement problem, where the application of quantum theory to the measuring devices themselves does not appear to lead to outcomes with definite values. The solution, first suggested by Hugh Everett in 1957, was the radical idea that all the possible outcomes do, in fact, occur (possible according to the predictions of quantum theory, as opposed to logically possible or conceivably possible). The wavefunction does not collapse. Instead, when the measuring apparatus seems to ascribe no definite value to the outcome, its quantum state is interpreted as describing several noninteracting copies of the system, each one exhibiting one of the different outcomes. When another system observes this device, the second system also branches into noninteracting copies, each now observing the outcome on a single copy of the device.

This approach has become widely known as the Many Worlds interpretation of quantum theory. For a while, it encountered considerable technical difficulties in accounting for the probabilistic predictions of quantum theory, but recent developments have, according to advocates of this point of view, answered these problems. Despite its apparent fantastic explanation, it remains in some respects the most conservative point of view, as it adheres most closely to the mathematical formalism of quantum theory.

The second point of view approaches the measurement problem by suggesting that there must be additional properties in the world. When the measuring device is described by a quantum state that does not give a definite value to the outcome of the experiment, and yet the observed fact is that the measuring device does give one out of a number of definite values for a property, this point of view holds that there must be facts in the world in addition to those given definite values by quantum theory. These additional facts have been termed *hidden variables*.

The most developed interpretation of quantum theory following a hidden-variable approach is the de Broglie–Bohm theory, first propounded by de Broglie, and independently rediscovered by David Bohm in 1952. In its most basic form, it attributes to atomic objects a well-defined position, in addition to a quantum state. The quantum state evolves in the manner of a wave, giving the interference and wavelike properties, while the position is guided by the quantum state. In this way, it answers the wave–particle duality problem by suggesting, not *wave or particle* but *wave and particle*. The measurement problem is solved by the additional position variable picking out which outcome has actually occurred, avoiding the need to postulate all outcomes actually occurring. The de Broglie–Bohm theory shares with the Many Worlds theory the idea that the quantum state is a representation of real physical properties. Advocates of Many Worlds have argued the de Broglie–Bohm interpretation is therefore also committed to the belief in the physical existence of the other worlds.

An alternate, if less well developed, approach to hidden variables asserts that only the hidden variables themselves, and not the quantum state, represent physically real properties. This is often

called “ Ψ -epistemic,” as it regards the quantum state (traditionally represented by the Greek character Ψ [psi]) as representing only epistemic uncertainty as to the actual state of the world. What distinguishes this from empiricist approaches is the idea that not only are there objective properties to the world, but that these properties can be meaningfully discussed and analyzed. The empiricist regards the quantum state as epistemic uncertainty about future measurement outcomes. The Ψ -epistemicist regards the quantum state as epistemic uncertainty about the state of the world.

The third point of view solves the measurement problem by proposing that quantum theory applies well to atomic objects, but that the dynamics of quantum theory must be modified to apply to large objects consisting of a great many atoms. The collapse of the wavefunction occurs through a physical process that differs from the normal quantum evolution. This is often known as an objective, or dynamical, collapse.

This novel process must be sufficiently slow that atomic objects still display quantum interference effects on the timescales that have been observed, while aggregate objects lose quantum interference effects and become essentially classical by the time they are directly observable. The two most developed proposals for this are the Ghirardi-Rimini-Weber and continuous spontaneous localization models. A characteristic feature of these approaches is that they must make experimental predictions that deviate from quantum theory at some point.

Owen J. E. Maroney

See also Kinetic (Molecular) Theory; Nanotechnology; Scientific Realism; Scientific Revolutions; Physical Theory

Further Readings

- Albert, D. (1994). *Quantum mechanics and experience*. Harvard University Press.
- Bacciagaluppi, G., & Valentini, A. (2013). *Quantum theory at the crossroads: Reconsidering the 1927 Solvay conference*. Cambridge University Press.
- Bassi, A., & Ghirardi, G.-C. (2003). Dynamical reduction models. *Physics Reports*, 379, 5–6.

- Holland, P. (1995). *The quantum theory of motion*. Cambridge University Press.
- Kuhn, T. (1987). *Black-body theory and the quantum discontinuity, 1894–1912*. University of Chicago Press.
- Rae, A. (2012). *Quantum physics: Illusion or reality?* Cambridge University Press.
- Saunders, S., Barrett, J., Kent, A., & Wallace, D. (Eds.). (2012). *Many worlds? Everett, quantum theory, and reality*. Oxford University Press.
- Timpson, C. (2008). Quantum Bayesianism: A study. *Studies in History and Philosophy of Modern Physics*, 39(3), 579–606.

PHYSICS, RENAISSANCE

Historians generally date the period of European history known as the *Renaissance* (French for *rebirth*) from the mid-14th century C.E. through the 15th and 16th centuries, beginning in Italy with the revival of learning from classical antiquity, and later spreading throughout Europe, aided by the invention of the movable-type printing press by the German craftsman Johannes Gutenberg around 1440. The Republic of Florence, Italy, is commonly thought to be the birthplace of the Renaissance, owing to its political structure, its civic and social peculiarities, the patronage of the Medici (its dominant family) and the migration of Greek scholars and their texts to Florence following the Ottoman Turks’ conquest of Constantinople, the capital of the Byzantine Empire, in 1453.

The Renaissance was a time of great social and cultural change that was felt in politics and religion, in art and architecture, music, philosophy, and the natural sciences. In the literature, authors began increasingly to write not in Latin but in native (vernacular) European languages. Renaissance artists studied how to depict objects in the same way they are perceived by the human eye, according to principles of perspective. Voyages of exploration over land and sea led to greater contact with other cultures and the spread of ideas. During the Middle Ages, understanding of the physical world had relied on beliefs in the supernatural, and it was thought the nature of the universe could not be explained by systematic observation and

experimentation. Renaissance thinkers changed this view with the radical concept that the universe could be studied and approached objectively. Mathematicians discovered that certain mathematical relationships were embodied in the natural world. The era's intellectual transformation acted as a bridge between the Middle Ages and the Modern Era. This entry focuses on the major developments in the domain of physics and related disciplines based on direct observation, such as astronomy and the study of magnetism; and the work of such key figures as Nicolaus Copernicus, Galileo Galilei, Johannes Kepler, and William Gilbert.

Astronomy and Physics

During the Renaissance period, the fields of astronomy and physics shared much common ground, and many theories by physicists sought to explain cosmological phenomena.

In modern science, there is still much overlap between astronomy and physics, as physicists use the universe as their *laboratory* for testing theories through observation. Astrophysicists and cosmologists study phenomena such as the effects of gravitational lensing and use increasingly powerful telescopes in an ongoing effort to understand the origins and nature of the universe. These observational practices were pursued without the aid of optical instruments by—most notably—the ancient Greeks, Babylonians, Aztecs, and Mayas, among others, and have been refined through the course of history, including the Renaissance period, when scholars such as Nicolaus Copernicus and Galileo Galilei used astronomical observations to propose theories to explain the universe's principles of motion.

Copernicus

Nicolaus Copernicus (Mikolaj Kopernik), who lived from 1473 to 1543, blended cosmology, mathematics, and physics in an effort to design a model of the universe. Born in Poland to a family of wealthy merchants, he studied liberal arts at the University of Krakow and traveled to Italy for further education preparing for a future in canon law. He studied astrology there as well as astronomy. Today, of course, astronomy is taken much more seriously in the scientific community than is astrology.

While attending the University of Bologna, Copernicus worked and lived with astronomy professor Domenico Maria de Novara, whom he helped make observations of the heavens. When he returned to Poland after graduation, one of the rooms in a large tower in his town held an observatory, which gave him opportunity to study the night sky. Copernicus never became a priest, but he continued to work as a cleric.

In 1514, Copernicus created a handwritten book for his friends that stated his views on the universe. In it, he argued that earth was not the center of the universe, but that the sun was very close to the center. He also stated that the planet's rotation was responsible for the movement of the sun and stars, as well as changes in the seasons.

Copernicus proposed several revolutionary ideas about around the structure of the solar system and where the earth was positioned in the heavens. Copernicus's ideas were based on astronomical observations gathered over several millennia. What had hitherto been lacking was a coherent theory explaining why certain objects acted in particular ways.

Copernicus proposed that the planets revolve around the sun, and he, like the Greek philosopher, mathematician, and geographer Ptolemy, who had lived and worked in Alexandria during the 2nd century C.E., believed that their orbits were circular and required epicycles. Copernicus's ideas took time to be accepted, because they conflicted with Church doctrine, based on biblical scripture, that the earth was the center of the universe. His views still assumed, however, that the universe reflected a divine, harmonious model composed of perfect circles and spheres, without disorder or eccentricity. During the Renaissance, God's perfection was a central tenet of orthodox belief; to doubt it constituted heresy, and this would not begin to change until Galileo's struggles with Church doctrine during the following century.

At the age of 70, in the year of his death (1543), Copernicus published *De Revolutionibus Orbium Coelestium* (*On the Revolutions of the Heavenly Spheres*). In this book, he finally laid out his theory that the planets orbited the sun rather than the earth, as well as his model of the solar system and the path of the planets, which today we call the heliocentric or Copernican model.

Galileo

Galileo Galilei was born on February 15, 1564, in Pisa, Italy. He was the first of six children born to his father, Vincenzo, a well-known music theorist and musician, and Giulia Ammannati. The family moved to Florence in 1574, and Galileo began attending school at the Camaldolese monastery in Vallombrosa.

In 1583, Galileo enrolled at the University of Pisa to study medicine. With his talent and intelligence, he became entranced with several subjects, including physics and mathematics. The university taught according to an Aristotelian view, which was sanctioned by the Roman Catholic Church, and initially Galileo adhered to this viewpoint. Owing to financial difficulties, however, he left the university before completing his degree.

After exiting the university, Galileo continued to study mathematics, and he supported himself with minor teaching positions. He also began his study of objects in motion, which would last two decades. He published *The Little Balance*, which earned him some attention with his descriptions of the hydrostatic principles of weighing small quantities. His work then gained him a teaching job at the University of Pisa, but his criticisms of Aristotle left him isolated among his peers, and in 1592, his contract with the University of Pisa was not renewed.

This would not be the end of Galileo's teaching career, however. He found a new position teaching astronomy, geometry, and mechanics at the University of Padua. Having been entrusted with the care of his younger brother Michelangelo following their father's death, this appointment came at a fortunate time. Galileo stayed at Padua for 18 years, giving lectures and attracting followers from across Europe. This increased his fame and his desire to fulfill what he saw as his life's mission.

Galileo began gathering evidence that supported Copernican theory and contradicted Church doctrine and the teachings of Aristotle. In 1612, he published his work *Discourse on Bodies in Water*, which refuted Aristotle's ideas of why objects float in water, arguing that was because of the weight of the object in relation to the water displaced, not because of its flat shape.

The next year, Galileo published his work on sunspots, in which he observed and refuted

Aristotle's work that argued the sun was perfect. The same year, he wrote a letter to a student, stating that Copernican theory was not in opposition to Biblical passages. His argument was that scripture was written from an earthly perspective and that science was unique and more accurate in its approach. The letter became public knowledge, and Galileo was labeled a heretic.

By 1616, Galileo was ordered by Church officials not to "hold, teach, or defend in any manner" the Copernican theory regarding the motion of the earth. Galileo obeyed this order a long seven years because of his devout faith and because of his social reputation.

In 1623, one of Galileo's friends (Cardinal Maffeo Barberini) became Pope Urban VIII. He allowed Galileo to resume his studies in astronomy and even publish this work, but only if it was written objectively and did not present the Copernican theory as an undeniable truth.

In 1632, Galileo published the *Dialogue Concerning the Two Chief World Systems*, a discussion among three people: one who argues against Copernicus's heliocentric theory of the universe, one who supports it, and one who is impartial. Though Galileo claimed *Dialogues* was neutral, it was clearly not. The advocate of Aristotelian belief did not give an intelligent argument, and it was clear where Galileo's beliefs stood.

When Galileo invented his telescope, it changed the idea of perfection, as well as showed that a Copernican model of the universe was (albeit incomplete) much more accurate than any other previous models. For the first time, people looking through Galileo's telescope could see that the universe had deep imperfections. Galileo discovered that the moon had mountains and that it looked similar to earth. He also saw that Jupiter had its own moons revolving around it, which astronomers everywhere were shocked to hear. Galileo's work in astronomy was highly controversial, his work in physics was even more so.

Galileo is still known for his groundbreaking work in mass and gravity, though the story of his dropping objects from the Leaning Tower of Pisa is likely a fabrication. Galileo's theories and experimentation showed that a larger or heavy object did not fall to the ground more quickly than a lighter or small object. His work was based on the study of earlier Greek and Islamic

mathematicians, but he gathered all of these data and expanded it, developing a theory to explain how the universe worked.

Galileo's additional contributions included those revolving around the idea of force. Aristotle had showed that the velocity of an object depended on the external forces acting upon it, and in this assumption he was correct. Aristotle made no distinction, however, between acceleration and velocity. He believed that objects moved due to his theory of the four elements (earth, wind, water, and fire) and that each element tried to return to its natural state. A rock, for its part, falls quickly because it is seeking to return to the earth; a feather, by contrast, falls slowly because it is also of the air. It remained to be demonstrated until after Galileo's lifetime that it is air resistance that causes a feather to fall more slowly than a rock; in a vacuum they will fall at similar speeds.

Galileo did not stop there. He expanded his theory to show that an object with no external force pushing or pulling it would maintain a constant velocity, but external forces could cause deceleration or acceleration. This *law of inertia*, so called, which stated that an object would move at a consistent velocity until another force acts upon it, is still accepted as a first principle of physics today. Much of the work Galileo performed depended on his observation of pendula, but he worked with falling objects as well. He would not simply take objects to the top of a tall structure and drop them (without a way to control the variable), because there was no good way to observe and time the fall of these objects.

Galileo built a series of ramps at various inclines that had a smooth channel down which he could roll perfectly smooth balls. In these experiments, he discovered that it was the distance the object fell, not the weight, which was most important. Objects accelerated the further they fell, but weight did not affect this, and all objects accelerated at the same speed.

These experiments were an excellent example of the scientific method. Galileo used deduction and induction to create his experiments and gather results and conclusions. This earned Galileo the reputation of the father of experimental science for many scholars, though this can, of course, be debated.

Reaction by the Church

The Church's reactions to Galileo's publications was quick, and he was summoned to Rome for a trial. The Inquisition lasted about 10 months, and he was treated with respect and never imprisoned. In an attempt to break his spirit once and for all, however, Galileo was threatened with torture, and he was forced to admit that he had supported Copernican theory, but kept his theories that Copernicus's statements were correct to himself.

Galileo was convicted of heresy, which led him to spend the latter years of his life under house arrest. Though he was ordered not to have any visitors or have his any further works printed outside of Italy, he ignored both of these warnings. A French translation of his work was adopted, and a year later, copies of his work called *Dialogue* were printed in Holland. While under house arrest, Galileo did not remain idle. Instead, he wrote a summary of his life's work on the strength of materials and science of motion. This work was called *Two New Sciences* and was published in Holland in 1683. Galileo was ill and blind by this time, and he eventually died after suffering from a fever and heart palpitations. The Church, at the time, still did not accept his work.

Galileo's Legacy

In 1758, the Church lifted the ban on many of the works that supported Copernican theory, and in 1835, it accepted that heliocentrism was a correct model, altogether. By the 1900s, many successive popes had acknowledged Galileo's great works, and in 1992, Pope John Paul II expressed his regret about how the Church had handled Galileo's work and treated him personally.

Galileo made major contributions to our understanding of the universe, and his methods of making his discoveries are just as important as his findings. Given his significant role in the scientific revolution, he is often called the *father of modern science*.

Kepler

Johannes Kepler, who lived from 1571 to 1630, was a German astronomer who discovered the

three major laws of planetary motion. They are as follows:

1. The planets move in elliptical orbits with the sun at one focus.
2. The time taken to traverse any arc of a planetary orbit is proportional to the area of the sector between the arc and the central body (*area law*).
3. There is an exact relationship between the cubes of the radii of planets' orbits and the planets' periodic times (the *harmonic law*).

Kepler himself did not refer to these findings as *laws*. That would happen when Isaac Newton later derived them from a new, unique set of general physical principles. Newton believed that these laws were celestial harmonies that reflected God's greater design for the universe.

Kepler's discovery differed from Copernicus's heliocentric model in that it showed a dynamic universe, with the sun actively pushing the planets in noncircular orbits. Kepler's physical astronomy theories created a new problem for other important 17th-century thinkers to solve—including Newton.

In addition to his work with the solar system, Kepler made additional efforts to provide a new account of how vision occurs. He developed a unique explanation for how light behaves in the newly invented telescope, and he discovered several new, semi-regular polyhedrons. He even offered a new theoretical foundation for astrology, while restricting its ability to be taken as fact.

Kepler studied many different fields, and his unique approach inspired those after him to make further discoveries, reflecting the efforts of scholars in different disciplines working in tandem. Though Kepler was most interested in studying astronomy, during his time this field was taught as a mixed science that consisted of a physical and mathematical component, having elements in common with other studies, such as optics and music. It was divided into practical and theoretical categories, including theory and practical construction of planetary instruments and tables.

In the 16th century, there was no true *scientific community* to speak of. Schooling in Germany (where Kepler lived) and throughout the entirety of Europe was controlled by the Church (whether Catholic or Protestant). Local rulers used these

educational centers to consolidate the loyalty of their populations. One way they did so was by giving out scholarships to poor boys, who, feeling fortunate and blessed to receive an education, would be loyal to the local ruler throughout their lives.

Kepler was from a modest family in the town of Weil de Stadt, and he was one of the beneficiaries of these ducal scholarships. He attended the Lutheran *Stift*, or seminary, at the University of Tübingen, where he began to study in 1589. Boys who attended these schools were expected to become state functionaries, ministers, or schoolteachers. Kepler had planned on becoming a theologian.

Kepler's life did not turn out quite as he expected it would. He believed that Divine Providence led him to study the stars, and he believed that this was a religious vocation. Kepler studied under mathematics professor Michael Maestlin, one of the most advanced astronomers in his country. A former Lutheran pastor, he was one of the few people who adhered to the Copernican theory of the late 16th century, but he was shy about saying so in person or in print.

Maestlin lent Kepler his annotated copy of *De revolutionibus orbium coelestium libri VI* (*Six Books Concerning the Revolutions of the Heavenly Orbs*), written by Copernicus. Kepler found himself extremely attracted to this work, and Maestlin tutored him on its finer details. Kepler believed Copernicus's work was a revelation; it clearly showed the mark of divine planning. Kepler made it his mission to demonstrate what Copernicus had merely guessed, and he presented his explanations in philosophical and religious vocabulary.

Kepler was not the only scientist to believe that the earth/nature had a spiritual connection. His ideas differed, however, in the intensity with which he believed them. For example, he was very strongly attracted to the idea that the Christian Trinity was symbolized by a geometric sphere, and therefore, the visible, created world was a reflection of God's divine mystery. One of Kepler's favorite Biblical passages was as follows: "And the Word became flesh and lived among us" (John 1:14).

For Kepler, this meant that the divine archetypes were made visible as straight and curved forms that became the spatial arrangement of tangible corporeal entities. Kepler believed that God was a creative, dynamic being whose presence was symbolized by the sun, and whose force

occasionally moved the planet. For him, the natural world was a mirror for divine ideas. Kepler believed that the human mind was created to understand the world's structure, and he believed that the soul was of the utmost importance to one's being.

Kepler's ideas were presented in his first work, *Mysterium cosmographicum* (1596; "Cosmographic Mystery"), and he continued to work with these same ideas and issues throughout his entire life. During 1595, Kepler had a moment of illumination while teaching a small class in Austria. He realized that the spacing between the six Copernican planets could be explained by inscribing and circumscribing each orbit with one of the five regular polyhedrons. Kepler recognized Euclid's proof that there can only be five such mathematical objects composed of congruent faces, and he decided that this self-sufficiency must come with a perfect idea. He thought that if the ratios of the mean orbital distances were in agreement with the ratios obtained from inscribing and circumscribing the polyhedrons, then he could discover the architecture of the universe. Kepler did find agreement within 5%; his exception was Jupiter.

Magnetism

Gilbert

William Gilbert, who lived from 1544 to 1603, was an English physician who also studied electricity and magnetism, though he did not fully understand the complexity or exact principles of either of these fields. Gilbert studied at Cambridge University and became a renowned doctor shortly afterwards. At university, Gilbert studied compasses and magnets, which were essential in England, a seafaring nation.

Magnets, which are known for their properties of attraction and are also known as lodestones, were well known to the Greeks and Muslims, as well as sailors across the globe. However, none of these groups understood the exact reasons that compasses work. Many scholars thought that they were simply attracted to the North Star or a collection of magnetic mountains somewhere within the Arctic Circle.

William Gilbert was not satisfied with this type of reasoning, and he set out to do a number of experiments on magnets. He came to understand some of the properties of magnets by using a large, spherical lodestone and a compass needle that

could rotate freely. Gilbert found that magnets lose their power of attraction when exposed to high temperatures; that they often produce circular motions; and that rubbing certain metals with magnets also turns them magnetic.

Gilbert realized that magnetism and rotation were often linked, which led him to suggest that the earth was one large magnet—an idea that he explained in his work *De Magnete* in 1600. In this work, Gilbert also rejected the idea that the earth was central to the universe and instead proposed that the earth was a giant magnet with north and south poles that was part of a heliocentric model.

Gilbert's work was thought by his contemporaries to be a breakthrough, and it would go on to influence Galileo's views of the structure of the universe. His experimental method was also revolutionary, in that it may have been the first example of an experimental scientific method. In addition, Gilbert's work on a concept he called *electricity* would create a foundation for scientific thinkers long afterward, but especially during the Enlightenment.

Katie Moss

See also Mathematics, Renaissance; Physics, 19th Century; Physics, 20th Century; Physics, Antiquity; Physics, Contemporary; Physics, Enlightenment; Physics, Middle Ages

Further Readings

- Crombie, A. C. (1959). *Medieval and early modern science* (Vols. 1 and 2). Harvard University Press.
- Field, J. V. (1988). *Kepler's geometrical cosmology*. University of Chicago Press.
- Repcheck, J. (2008). *Copernicus' secret: How the scientific revolution began*. Simon & Schuster.
- Reston, J. (1994). *Galileo: A life*. Beard Books.
- Wilson, R. (2000). William Gilbert: The first palaeomagnetist. *Astronomy and Geophysics*, 41(3), 3.16–3.19.

PHYSICS, SOLID-STATE

Solid-state physics is defined as the study of solids of rigid matter, through various methods such as crystallography, metallurgy, electromagnetism,

and quantum mechanics. Solid-state physics explains the different properties of all solid materials and the properties of a collection of atomic nuclei and electrons interacting with electrostatic forces among themselves. It also typically formulates the basic and fundamental laws that govern the typical behavior of solids.

Since the 1960s, this branch of physics has been subsumed by more general category of *condensed matter physics*. This was done mostly to include fluids, which also in many cases readily share the same concepts and analytical techniques as those of solids. Thus, solid-state physics is now commonly discussed under the more comprehensive term, *condensed matter physics*, of which it is generally considered to be the largest branch. Condensed matter physics is actually the field of physics which deals with both the microscopic and macroscopic physical properties that matters possess. In particular, it is chiefly concerned with the extensive “condensed” phases that, in most cases, appear whenever the probable number of participating constituents in a particular system is quite large and the typical interactions between these constituents are fairly strong. The most familiar and examples of the condensed phases are of course solids and liquids, which typically arise from the strong electromagnetic forces that exist between atoms. Solid-state physics is concerned with studies that explain how large-scale properties of relatively solid materials actually result from their intrinsic atomic-scale properties. Thus, it can be estimated that solid-state physics forms the appropriate theoretical basis of the vast application-based field of *materials science*. It also has some direct applications, as for example in relation to the hardcore technologies concerning semiconductors and transistors. Solid materials, as we know, are typically formed from densely packed atoms. These atoms interact intensely among themselves. These sophisticated interactions mostly produce the specific mechanical (e.g., elasticity and hardness), magnetic, thermal, optical, and electrical properties of solids. Depending on the type and kind of material involved and the relative conditions within which it was actually formed, the atoms may constitutively be arranged in a specific regular, typical geometric pattern (like crystalline solids, which mainly include all the metals and ordinary water or ice) or quite

irregularly (e.g., a grossly amorphous solid such as common window glass).

Crystals and Crystallography

The bulk of solid-state physics, viewed mostly as general theory, is extensively focused on crystals. Crystalline materials are solids with a highly organized atomic structure typically based on a regular repeated pattern. Most solids that are known happen to be crystalline. There are also some known liquid crystalline colloids of the nanoparticles. Much more progress has been effectively made in first decades of the 21st century to understand the behavior of these crystalline solids rather than that of the noncrystalline materials, since most of the concerned calculations are easier to perform on crystalline materials. Understanding the basic electrical properties of these solids is of pivotal importance to modern-day society and advancing technology.

Primarily, this phenomenon is because of the periodicity exhibited by the atoms in a crystal—its defining characteristic, which greatly facilitates mathematical modeling. Similarly, crystalline materials quite often have magnetic, electrical, mechanical, or optical properties which can be exploited for the purposes of engineering. The typical forces that exist between the constituent atoms in a crystal can take a variety of structural forms. For example, a crystal of sodium chloride (i.e., common salt) is basically made up of sodium and chloride ions, which are held together with many ionic bonds. Sodium chloride specifically crystallizes in the form of a cubic lattice, but with a very different unit cell. The structure of sodium chloride consists of exactly equal numbers of both sodium and chlorine ions, stringently placed at every alternate point of a very simple cubic lattice. In contrast, the noble gases (helium, neon, argon, krypton, xenon, and the radioactive radon) never undergo any of these typical types of bonding. In solid form, all the noble gases are actually held together with the Van der Waals forces acting between them, resulting from the total polarization of the surrounding electronic charge cloud on each of these atoms. The actual difference that exists between the types of solid typically result from the fundamental differences between their intimate bonding. Many properties of these solid

materials are vastly affected by their crystal structure. This structure can be effectively investigated using a wide range of various crystallographic techniques including neutron diffraction, X-ray crystallography, and electron diffraction. Various diffraction techniques solely exploit the phenomenon of scattering of radiation from a large number of available sites. There are various diffraction techniques which are currently being employed that result in typical diffraction patterns. These typical patterns are gross records of the various diffracted beams produced. The sizes of individual crystals found inside a crystalline solid material do vary extensively depending on the type of material involved and the prevalent conditions when the material was formed. Most of the crystalline materials encountered in everyday life are by nature polycrystalline, with the individual crystals being microscopic in scale, but macroscopic single crystals can also be very easily produced either by natural (e.g., diamonds) or artificial methods.

A single crystal has a typical periodic atomic structure almost across its entire volume. At long-range typical-length scales, nearly every atom is generally related to almost every other grossly equivalent atom in the utmost structure by rotational or translational symmetry. These polycrystalline materials are mostly made up of clusters of many tiny single crystals (called *grains* or *crystallites*). Polycrystalline materials have a high degree of restricted order over many molecular or atomic dimensions. Grains (*domains*) are typically separated by the grain boundaries. Their atomic order can greatly vary from one particular domain to the next. The grains are usually 100 nm to 100 μ m in diameter; polycrystals which have grains that are less than 10 nm in diameter are actually nanocrystalline. On the other hand, amorphous or noncrystalline solids are made up of randomly orientated molecules, atoms, or ions that do not necessarily form exactly defined lattice structures or patterns. Amorphous materials range only within a few molecular or atomic dimensions; such materials do not actually have any particular long-range order, but they seem to exhibit varying degrees of short-range order structures. Examples of amorphous materials include the amorphous plastics, silicon, and glass. These amorphous

silicates are readily used in thin-film transistors and solar cells.

Crystallography is a specific branch of science which deals with the basic geometric description of all crystals and the internal arrangement of their atoms. The symmetry of a typical crystal is obviously important for its profound influence on the crystal's general geometric properties. These structures should hence be effectively classified into different types according to the basic symmetries they possess. All energy bands can be distinctly calculated when their structure has well been determined. In the case of elaborate crystallography, when studying a crystal lattice the main interest is in the geometrical properties of the crystal; hence one may suppose replacing each of the atoms by a specific geometrical point located exactly at the typical equilibrium position of that particular atom. Relatedly, there exists an infinite array of points in the space; each point has more or less identical surroundings to all the other points. All these arrays are regularly arranged in a periodic manner. Crystal structures can be obtained by attaching atoms, groups of atoms, or molecules which are called basis (motif) to the lattice sides of the lattice point. The unit cell (group of atoms, ions, or molecules) can be considered the smallest component of the crystal, which when stacked together with pure translational repetition reproduces the whole crystal. Real crystals actually feature irregularities or defects in their ideal arrangements, and it is mostly these defects that critically determine almost all of the mechanical and electrical properties of the real materials.

The crystal lattice has the ability to vibrate. These intense vibrations are typically found to be super quantized; these exceptionally quantized different vibrational modes are known as *phonons*. A phonon is actually a collective excitation which occurs in a periodic arrangement of molecules or atoms in a typical solid and some liquids. Although it is often referred to as a typical quasiparticle, the phonon actually represents a very excited state in the actual mechanical quantization of most of the different modes of vibration of typical elastic structures of rapidly interacting particles. Phonons typically play a major role in many of the physical properties of condensed matter, such as electrical conductivity and thermal conductivity. The study of these phonons is an

important part of contemporary condensed matter physics. The concept of these phonons was first introduced in 1932 by the Russian physicist Igor Tamm. The term *phonon* comes from a Greek word that means anything that translates sound or voice. The phonons of longer wavelength mostly give rise to sound, whereas the shorter wavelength and higher frequency phonons typically give rise to heat. In the insulating solids, phonons are also the main and primary mechanism by which gross heat conduction takes place. Phonons are by far also necessary for critically understanding the exact lattice heat capacity of any solid, as explained in the Einstein model of the solid and later clearly demonstrated by the Debye model, developed by Peter Debye in 1912, for estimating the phonon contribution to the heat capacity of a solid.

Properties of solid materials such as heat capacity and electrical conduction are largely investigated by solid-state physics. The Drude model, proposed in 1900 by the German physicist Paul Drude, was one of the earliest models explaining electrical conduction in solids, which mainly applied the kinetic theory to the free and bound electrons in a solid. By assuming that the solid material actually contains many immobile positive ions and a typical “electron gas” of classical, basically noninteracting electrons, the Drude model was able to clearly explain thermal and electrical conductivity, and the Hall effect in these metals, although the electronic heat capacity is greatly overestimated. Arnold Sommerfeld for the first time combined the classical Drude model with the concepts of quantum mechanics in his established free electron model of 1927 (more frequently called the *Drude–Sommerfeld model*). Here, the free electrons are actually modeled as a typical Fermi gas—in other words, a gas of particles that obey the critical quantum mechanical model relating to Fermi–Dirac statistics. This free electron model actually gave improved predictions about the heat capacity of the metals. The model is actually an almost total modification of the basic electron model that mainly includes a quite typical weak and periodic perturbation, which is basically meant to variably model the basic interaction between the exact conduction of these electrons and the typical ions in a basic crystalline solid. By introducing the new idea of electronic bands, the

theory properly explains the existence of conductors, semiconductors, and insulators. This almost free electron model typically rewrites the Schrödinger equation for the exact case of a periodic potential. The available solutions in this case are better known as *Bloch states* (after the Swiss physicist Felix Bloch). Bloch’s theorem applies only to exact periodic potentials, and these unceasing and random movements of atoms within a crystal disrupt the regular periodicity, but it has actually proved to be a highly valuable form of approximation, without the use of which most of the analysis of solid-state physics would be totally intractable.

Current Research Topics

Contemporary researchers are mainly focused on the following topics in relation to solid state physics.

Strongly Correlated Materials

These are a broad class of electronic materials which often show unusual (technologically very useful) magnetic and electronic properties, such as half-metallicity or metal–insulator transitions. The most essential feature that clearly defines clearly these exquisite materials is that the behavior of their energized electrons can never be described almost effectively in typical terms of viable noninteracting entities. The major theoretical models of the related electronic structure of these strongly correlated materials must sincerely include electronic correlation to be stringently accurate.

Quasicrystal or Quasiperiodic Crystal

These structures are typically ordered but not at all periodic. A quasi-crystalline pattern most obviously can fill all available space continuously, but it severely lacks translational symmetry. Although the crystals, according to the general classical crystallographic restriction theorem, typically possess only two-, three-, four-, or six-fold rotational symmetries, herein the Bragg’s diffraction pattern of these quasicrystals shows sharp peaks with most other symmetry orders, for instance, about five-fold.

Spin Glass

Spin glass is actually a highly disordered magnet, wherein the typical magnetic spins of the individual component atoms (the relative orientation of the magnetic poles; i.e., north and south poles in the exact three-dimensional space) are not at all aligned in any regular pattern. The term *glass* actually comes from a typical analogy between the positional disorder of a mostly conventional, chemical glass (e.g., window glass), and the highly magnetic disorder in a spin glass. In window glass or in case of any amorphous solid, the typical atomic bond structure is highly irregular; whereas in contrast to this phenomenon, a crystal has a very uniform pattern of regular atomic bonds. In all ferromagnetic solids, all magnetic spins align rightly in the same direction; this would be typically analogous to a crystal.

High-Temperature Superconductivity (HTS or High- T_c)

These are typical solid materials that behave as efficient superconductors mostly at unusually high temperatures. The first ever HTS was actually discovered in 1986 by two IBM researchers, K. Alex Müller and Georg Bednorz, who were later awarded the 1987 Nobel Prize in Physics for this discovery. Their work led to an important breakthrough in the area of discovering superconductivity in the ceramic materials.

Syed Feroj Ahmed

See also Physics, Gravitation Theory; Physics, Quantum Theory

Further Readings

- Garbovskiy, Y. A., & Glushchenko, A. V. (2011). "Liquid crystalline colloids of nanoparticles: Preparation, properties and application." In R. Camley & R. Stamps (Eds.), *Solid state physics*. Elsevier.
- Gránásy, L., Pusztai, T., Warren, J. A., & Douglas, J. F. (2004). A general mechanism of polycrystalline growth. *Nature Materials*, 3(9), 645–650.
- Levine, D., & Steinhardt, P. J. (1984). Quasicrystals—A new class of ordered structures. *Physical Review Letters*, 53.
- Slutsky, L. (1960). Solid state physics. Advances in research and applications. Volume 10. *Journal of the American Chemical Society*, 82(23), 5961–6211.

Tan, Y. W., Zhang, Y., Bolotin, K., Zhao, Y., Adam, S., et al. (2007, December 14). Measurement of scattering rate and minimum conductivity in graphene. *Physical Review Letters*, 99, 246803.

PHYSICS, THERMODYNAMICS

Thermodynamics is the study of the relationship between temperature, heat, and bulk properties of matter such as volume, pressure, and mass. It developed in the 19th century through attempts to understand the general principles behind the newly developing technology of steam engines, and the failures of earlier theories of heat. Material bodies are treated at an aggregate level, parameterized by a few operationally well-defined properties, without reliance on detailed underlying models of how matter is composed. A few general principles, combined with the results of characteristic measurements for a given material, allow detailed conclusions to be drawn about the thermal properties and behavior of the material and its interactions with other materials.

In the late 19th and early 20th centuries, thermodynamics was supplanted, as a fundamental theory of physics, by statistical mechanics, which seeks to account for the same phenomena in terms of the behavior of large numbers of atoms and molecules. Statistical mechanics continues to use theoretical terms that originated in the theory of thermodynamics, and the relationship between the two theories is often considered the paradigmatic example of theory reduction in the sciences. After tracing the historical development of thermodynamics, this entry presents the core concepts of the study, outlines the four laws used to define thermodynamic properties and behavior, and concludes with an account of the developments in physics since the 1860s which have led to revision and refinement of the conceptual foundations on which thermodynamics was based.

Historical Development

The beginnings of thermodynamics lay in the development of an operational definition of temperature. Although primitive notions of *hotter than* and *colder than* are evident from the senses,

the development of a numerical temperature scale required materials that would consistently expand as they were heated and two fixed, reproducible temperatures that could be used to calibrate thermometers. By the middle of the 18th century, the boiling and freezing points of water had started to become adopted to calibrate thermometers based upon the expansion of mercury or of gases.

Once thermometry had advanced, comparisons could start to be made as to the extent to which temperatures changed as different kinds of actions were performed upon different materials. Joseph Black's observation that ice stayed at the same temperature, while continuing to absorb heat from other bodies during melting, led to the idea of heat or cold as a substance called *caloric*. Bodies with a lot of caloric would be hot. Bringing them into thermal contact with cooler systems would result in a flow of caloric from one to the other, until they reached the same temperature. By measuring the change in temperature brought about by this loss of caloric (calibrated by means such as the largest quantity of ice that could be melted by absorbing the caloric), the different responses of different materials to caloric could be measured.

In the first half of the 19th century, a series of experiments cast doubt upon the caloric theory of heat and led to the development of thermodynamics. In 1798, Count Rumford showed that the friction involved in boring a cannon could seemingly produce unlimited amounts of heat, a problem if heat was a conserved substance. Julius von Mayer, in 1842, and James Joule, in 1843, showed there was a fixed amount of heat that would be produced by a fixed amount of mechanical work. The quantification of the mechanical equivalent of heat suggested that heat and work were interconvertible. Instead of two distinct things, they were different instances of a single conserved property: energy.

The development of steam engines in the 18th century showed that the flow of heat could be used to extract mechanical work, such as driving a piston or lifting a weight, but even the best steam engines seemed to produce large amounts of excess heat, compared with the work extracted. This problem was first effectively studied by Sadi Carnot in 1824. Although working with the caloric theory, he discovered a limitation on the extent to which the flow of heat between bodies at

different temperatures could be used to extract mechanical work, and the limitation depended only on the difference in temperature between the hottest and coldest bodies present.

Émile Clapeyron removed the dependency on the caloric theory from Carnot's work, in 1834, but it was Rudolf Clausius, in 1851, and William Thomson, in 1852, who combined Carnot's work with the conservation of energy to produce what would become the theory of thermodynamics (Thomson coined the term in 1854). The most significant breakthrough was when Clausius showed that Carnot's analysis implied the existence of a new property, which could increase but never decrease, and which gave a measure of the limitations on which heat energy could be converted back into mechanical work. In 1865, Clausius named this quantity as *entropy*.

The theory of thermodynamics came to be expressed in two laws: the conservation of energy and the impossibility of entropy decreasing. By the 1930s, two further laws were added, clarifying the properties needed to define a temperature scale, and established how properties of materials behaved as the temperature approached its minimum possible value.

Almost as soon as the theory of thermodynamics had been developed, however, it was being replaced. From 1857 onward, Rudolf Clausius and James Clerk Maxwell started to develop an explanation of thermodynamic properties based upon atomic theories of matter; this was further developed by Ludwig Boltzmann, beginning in 1872, and Josiah Willard Gibbs up to 1903. Although these explanations faced significant conceptual challenges, their great success in accounting for the specific form of empirically measured thermodynamic properties led to its being widely accepted by the early 20th century.

Core Thermodynamic Concepts

The core concepts in thermodynamics are *mechanical work*, *thermal equilibrium*, *temperature*, *heat*, and *entropy*.

The notion of mechanical work in thermodynamics is an expression of the energy given to, or extracted from, a system by means of changes in controllable parameters, such as the height or a weight or the volume of a gas. It is derived from

Newtonian mechanics, where a force acting on a system as the system moves through a fixed distance increases the energy of the system by the product of the force and distance. The standardized measure of work is the distance that a fixed mass can be moved through a gravitational field. In a gas, the notion of force is generalized to pressure, and distance is generalized to volume. When the gas changes volume at a constant pressure, the energy of the system changes by the product of the pressure and the change in volume. If the volume decreases, the energy must increase, and work must be performed upon the gas. If the volume increases, then the energy can go down, and the pressure can be used to extract work from the expanding gas. This is how the expansion of a gas against a piston can be used to draw energy out of a system. In general, the mechanical work required is equal to the product of the change in a parameter and a generalized force acting upon that parameter.

Thermal equilibrium plays a central role in the theory. A system is in equilibrium if it is not changing over time. For thermodynamics, a more complicated notion is needed for thermal equilibrium. If a system is thermally isolated, so no heat can be exchanged with other systems, and all the controllable parameters are fixed at some value, then according to thermodynamics the system will approach coming to rest in a unique equilibrium state. This state can be characterized as a function of the controllable parameters, the temperature and the energy of the system. This function, the *equation of state*, is the starting point of most thermodynamic analyses. For each material, it must be empirically determined which are the relevant controllable bulk parameters and what form is taken by the equation of state. When two systems are brought into thermal contact with each other, they will exchange heat until they reach a joint state of thermal equilibrium with each other.

The notion of temperature needed careful clarification for its use in thermodynamics. A thermometer is, at its simplest, a system which has an observable property that increases steadily and consistently as the system is heated. A thermometer has no other controllable parameters, so its equation of state depends only on the property and energy. If a thermometer is brought into thermal

contact with another system, heat will be exchanged until they have reached thermal equilibrium. The observable property of the thermometer then indicates the temperature of the system. To be a good thermometer, the temperature of the thermometer would need to be very sensitive to the changes in its internal energy, so that only a small quantity of energy needs to be exchanged with the measured system, introducing negligible changes.

However, this might still allow a system to show different temperatures depending upon the type of thermometer used. The key idea for thermodynamic temperature is that there is a universal temperature scale, so an ideal thermometer is measuring an actual property of the world. For this, thermodynamics requires that whenever two systems are individually in thermal equilibrium with a third system, they must also be in thermal equilibrium with each other. If the third system is the thermometer, this ensures that it records the same temperature for all systems in thermal equilibrium with each other. If two of the systems are thermometers, the requirement ensures that the different thermometers measure the same property, even if they are constructed from different materials and physical properties. Different thermometers might define different scales, but a conversion scheme in principle exists between the different scales. The Fahrenheit and Celsius scales are two examples. The Kelvin scale was based on the equation of state of an ideal gas, which led to the understanding that there was a minimum point on all temperature scales: the absolute zero of the Kelvin scale.

When two systems are brought into thermal contact, energy can be exchanged between them without any work being performed. In thermodynamics, this kind of energy exchange is called *heating*. Prior to thermodynamics, heat was considered a kind of substance, like a fluid, and this language persists (e.g., heat *flows*). Now, however, only the total internal energy of a system is considered to be a property of a system in thermodynamics. The energy is not divided into separate quantities of heat and work; instead heat and work quantify two ways in which the internal energy of a system may be changed.

Probably the most important and unexpected concept in thermodynamics was the existence of

entropy. When an object's energy changes by a small amount, whether through work or heat, the change in its entropy is given by the ratio of the change in its energy to its temperature (measured in the Kelvin temperature scale). In any exchange of energy between systems, the second law of thermodynamics states that the sum of the entropies of all systems involved cannot go down: It must either stay constant or increase.

When two objects at different temperatures are brought into direct thermal contact, such as an ice cube placed in a glass of water, they exchange heat until they come to the same temperature. Conservation of energy requires that the heat emitted by one is equal to the heat absorbed by the other. The object losing heat decreases in entropy, while the object absorbing heat increases in entropy. As the change in entropy is given by the same quantity of heat but divided by their different temperatures, the hotter water has a smaller change in entropy than the colder ice. For the total entropy to increase, the ice needs to increase entropy and the water needs to decrease entropy, so heat must be emitted by the water and absorbed by the ice. Ice in water melts, cooling down the water. This holds in general: When two objects at different temperatures are brought into thermal contact, heat must be emitted by the hotter object and absorbed by the colder. Entropy not only unites the disparate phenomena of heat spontaneously flowing only from hotter to colder objects but provides an explanation of why no examples of the reverse processes can be found.

The notion of entropy explains much more than this. It was developed from the study of heat engines. A steam engine allows heat to flow from a heated body (steam) into a lower temperature system (the heat sink, typically the environment) but extracts some of the energy emitted by the hotter body and converts it to mechanical work. In the case of a steam engine, this is typically by driving a turbine, which might pull a pulley, drive a pump, or generate electricity. Carnot also contemplated a heat pump, the reverse of a heat engine, in which heat is extracted from a cold system (making it colder) and transferred to a hotter system. Such a process would require work to be performed, adding to the heat in the system. Modern examples of heat pumps are refrigerators and air conditioners. If a heat pump and a heat

engine work together, Carnot showed there is a maximum efficiency at which they can operate. In the case of the heat engine, there is a maximum amount of heat which can be extracted from the hot body and converted to work. For the heat pump, there is a minimum amount of work required to cool down the colder system.

In an entropic analysis of the heat engine, the extraction of some of the heat, in the form of mechanical work, reduces the amount of heat that is being absorbed by the colder object. This must reduce the extent of its entropy increase. The maximum amount of work that can be extracted is just enough so that the increase in entropy of the colder body is no more than the entropy reduction of the hotter body. Extracting any more than this would lead to a net decrease in entropy, so real heat engines cannot extract more than this amount of work.

For a heat pump, if heat could be transferred from a colder body to a hotter body without assistance, then the reduction in entropy of the colder body would be greater than the increase in entropy of the hotter body. This would require the total entropy to decrease. If, however, some work is performed upon the system, adding extra heat to the hotter object, then the entropy increase in the hotter object becomes larger. The minimum amount of work needed is just enough so that the entropy increase of the hot body just matches the entropy reduction of the colder body, and the net change in entropy is zero. Real refrigerators require at least this amount of work to operate, and they dissipate this work as extra heat. Entropy thus characterizes which processes are possible and which are not and also gives quantitative values on the amount of work required to effect changes in systems.

The concept of entropy, and the law of its increase, unified a large number of phenomena, explaining why some are widely observed and their reverse processes are almost never observed.

The Laws of Thermodynamics

Thermodynamics is conventionally stated in terms of four laws (the exact phrasing of which varies from source to source) that are used to define thermodynamic properties and behavior. For

historical reasons, the laws are numbered from zero to three (the zeroth law, while logically prior to the first, second, and third laws, was only given a precise formulation by Ralph Fowler and Ernest Guggenheim in 1939). The first and second of these laws are often expressed as forbidding types of perpetual motion.

The zeroth law is intended to assert the existence of a universal temperature scale. A typical formulation is: "If two systems are each in thermal equilibrium with a third system, they are always in thermal equilibrium with each other." This ensures that different thermometers showing systems in thermal equilibrium with each other are at the same temperature as each other. Temperature is therefore defined as a quantity that parameterizes thermal equilibrium.

The first law is often expressed as the existence of a conserved quantity called *energy*. A typical formulation is "The change of energy of a system is equal to the sum of the heat absorbed by the system and the work done upon the system." Because heat and work characterize exchanges of energy between systems, the gain in energy by one system must equal the loss in energy by all other systems.

The first law is often stated as the impossibility of a *perpetual motion machine of the first kind*. A perpetual motion machine is one that is able to complete a sequence of operations in a fixed amount of time that return it to its original state, and it is therefore able to perform the cycle of operations again any number of times. A machine of the first kind is one that can perform such a cycle, in thermal isolation, and perform work upon another system while doing so. It would violate the first law of thermodynamics, because its internal energy must remain unchanged and the heat absorbed must be zero, but the work extracted from the system is not zero. This would mean energy was not conserved over the course of the cycle.

The second law expresses the existence of a property called *entropy*, which can increase but cannot decrease. A typical formulation is "Heat cannot spontaneously flow from a colder system to a hotter system." As the entropy change in a system is given by the heat flow divided by the temperature, a flow of heat against the temperature gradient would result in a small entropy increase in the hotter system and a large entropy

decrease in the colder system, giving a total decrease in entropy. Entropy decreases are possible in individual systems, but only in processes where at least as much entropy increase occurs in other systems.

The second law is often stated as the impossibility of a perpetual motion machine of *the second kind*. A machine of the second kind is one which can perform a cycle of operations, returning to its original state, while extracting work from a single system, cooling it down in the process. Such a machine does not violate the first law of thermodynamics, as the work extracted equals the reduction in energy of the cooling system. However, the entropy of the cooling system would be decreasing, and there would be no compensating entropy increase anywhere else.

The third law of thermodynamics concerns the inability to reach the absolute zero of the temperature scale. A typical formulation is "The entropy of a system approaches its minimum value as the temperature approaches zero." At absolute zero on the temperature scale, the entropy of the system becomes independent of the values of the controllable parameters in the equation of state. Cooling a system down involves varying these parameters in cycles to extract energy from the system. As the entropy approaches its minimum, the entropy change from varying the parameters becomes even smaller, in such a way that no finite number of cycles can reduce the entropy all the way to its minimum. As a result, it is impossible to reach absolute zero temperature in a finite number of cycles.

Further Developments

From the 1860s, the kinetic theory of gases identified heat as the motion of atoms and molecules. *Pressure* is the force exerted by a large number of particles colliding with the side of a container holding them, and the temperature of a gas is an expression of the total kinetic energy of its motion. In 1872, Ludwig Boltzmann identified a function of the microscopic particles which could fulfill the role of entropy. Although highly successful, this raised significant conceptual problems for thermodynamics, with particular problems for the second law.

In 1867, Maxwell imagined a tiny shutter between two boxes containing a gas, initially at the same temperature and pressure on both sides

of the shutter. If the shutter can be opened only when a molecule approaches from the left, and closed when a molecule approaches from the right, then gradually the gas will accumulate in the box on the right. This will lead to a pressure difference between the two boxes, and the gas can be allowed to flow back into the box on the left, driving a turbine to extract work on the way. This would be a perpetual motion machine of the second kind.

In 1876, Johann Joseph Loschmidt raised the problem that the laws of physics are symmetric in time. If a system can approach thermal equilibrium, then it must be possible that a system can depart from thermal equilibrium, undermining the whole concept. Equally problematically, if processes that increase entropy are possible, then processes that decrease entropy should also be possible.

In 1890, Henri Poincaré showed that a closed system governed by Newton's laws, if left isolated for long enough, must return arbitrarily close to its initial state. So if an ice cube were dropped in a glass of water, and the system was completely isolated, the ice cube would melt but at some point in the future it would have to spontaneously reform.

The resolution of these problems is now generally considered to be that reductions in entropy are possible, but under normal circumstances are so unlikely that they are never observed. The source of this implausibility is that the initial state of the universe was in an extremely low entropy state, and this biases the processes that occur against observing reductions of entropy on any reasonable timescale.

Owen J. E. Maroney

See also Kinetic (Molecular) Theory; Scientific Revolutions

Further Readings

- Fermi, E. (2000). *Thermodynamics*. Dover.
- Chang, H. (2007). *Inventing temperature*. Oxford University Press.
- Truesdell, C. (1980). *The tragicomical history of thermodynamics*. Springer.
- Uffink, J. (2001). Bluff your way in the second law of thermodynamics. *Studies in History and Philosophy of Modern Physics*, 32(3), 305–394. [https://doi.org/10.1016/S1355-2198\(01\)00016-8](https://doi.org/10.1016/S1355-2198(01)00016-8)

PHYSICS, 20TH CENTURY

In the 20th century, physics expanded at a rapid pace, on the one hand continuing to further clarify and articulate the theories it inherited from 19th-century physics, and on the other coming to grips with the various revolutionary developments arising at the turn of the 20th century. These developments were prompted by a number of observations that seemed simply incompatible with our basic understanding of the laws of nature—for example, black body radiation, radioactivity, and the mismatch between electromagnetism and mechanics. The 20th century saw important developments in the theoretical tools at the disposal of physicists in their efforts to interpret, understand, predict, and control the physical world.

This entry highlights several important conceptual developments in physics in the 20th century, including quantum theory, particle physics, the special theory of relativity, the general theory of relativity (along with the big bang theory). The entry also mentions briefly some other developments in theory flowing from 20th century physics, and their connections to other physical disciplines. Because of space limitations, we must leave details of these other important developments and technical issues to other entries in this encyclopedia.

Background

Constituents of Matter

Although in the west the Greeks first hypothesized the concept of the atom around two-and-half millennia ago, for long afterward there was no true theoretical development of the concept. During the 19th century C.E., the facts concerning the existence of atoms had been indirectly confirmed, and a more detailed theoretical picture was being developed in order to make use of the concept. The picture of the atom at that time was much different from what it is now. In the 19th century, the structure of atom was thought to be that of a spherical globule with more-or-less uniform density. Nowadays, the actual structure is thought to consist of a very small and heavy nucleus, around which swarm very lightweight,

negatively charged particles called *electrons*. The name *atom*, meaning “indivisible” in Greek, has thus been shown to be a misconception. The atom is not the basic building block or the smallest part of a molecule; rather, it is constructed of smaller units. Furthermore, the nucleus is composed of neutrons and protons, each of which actually weighs almost 2000 times more than an electron. As their names suggest, the neutron is neutral, with no amount of electric charge, whereas the proton is positively charged. It was 1897 when the existence of the electron was confirmed; previously, it had been thought that the electron was a form of liquid matter rather than a microscopic particle. In the first half of the 20th century it was commonly supposed by scientists that the basic components of matter were neutrons, protons, and electrons. Since then, high-energy accelerators have disclosed the fact that neutrons and protons are made up of three quarks. *Quarks* are point-like, microscopic units with electric charges that are only one-third and two-thirds of the total charge of the electron. In short, matter could be divided by the scientists into smaller units to an ever-increasing extent. After centuries of experiments and discoveries, a splendid vivid picture of the basic constituents of matter has been achieved. It remains to be seen when this “reductionism” of matter will cease.

Developments in Radioactivity

Owing to the evolutionary phase of physics during the course of the 19th century, highly complex problems could be handled by classical mechanics: detailed analysis of macroscopic situations; kinetic theory and thermodynamics were developed significantly; physical and geometrical optics in terms of electromagnetic waves could be well understood; there was widespread acceptance for the laws of conservation of momentum (and mass) and energy. The impact of these and other developments produced a general idea that all physics-related crucial laws had been discovered. The only aim of the discipline from then on would be to pinpoint and solve the minor issues of physics and develop better means of measurements and methods. But the notion of the completeness of the classical theories suffered a serious setback, and around 1900 critical doubts

arose about them. For example, inadequacies had begun to appear in James Clerk Maxwell’s theory regarding the photoelectric effect, black body radiation, and energy distribution. Paradoxes were noticed with some of the theories when pushed to their limits of application. Some prominent physicists, such as Wilhelm Wien, Emil Cohn, and Hendrik Lorentz, believed that all physical laws could be based on Maxwell’s equations when modified appropriately. However, it was soon realized that new ideas were required in order to resolve the problems of classical physics. Indeed, the world of physics was shaken by a major revolution at the beginning of the 20th century that marked the emergence of modern physics. For his crucial discoveries of electrons and isotopes, and the construction of the first mass spectrometer, the British physicist J. J. Thomson was awarded the Nobel Prize for Physics in 1906. Unexpected forms of radiation were detected by researchers in the 19th century. In 1895, the sensational discovery of X-rays took place through the work of William Röntgen. The emission of radiation by certain elements was discovered by Henri Becquerel in 1896. With the discovery by Marie and Pierre Curie of the new radioactive elements radium and polonium, questions began to be asked about the nature of matter and indestructibility of the atom. Two of Becquerel’s radiation forms were derived from the element helium, and electrons through the efforts of Frederick Soddy and Ernst Rutherford. Two types of radioactivity were identified by Ernst Rutherford; in 1911, he proposed that the atom consisted of negatively charged electrons surrounding a positively charged nucleus at the center. However, the stability of this structure was not in accordance with classical theory. In the late 19th century, two other theories arose that could not be explained by classical theory. One of these had to do with Albert A. Michelson and Edward W. Morley’s experiment that disconfirmed the existence of the so-called *luminiferous aether*, hitherto a preferred frame of reference for explaining electromagnetic phenomena. Throughout the 1930s, a preeminent focus of physical and chemical sciences was that of research on radioactive decay and radiation. It is during this time that the doors to the domain of atomic energy were opened through the discovery of nuclear fission.

Quantum Mechanics

Quantum mechanics is one of three important 20th-century revolutions in physics. It basically governs the dynamics of the microscopic and subatomic objects such as the atoms and electrons. Its development was meant to solve the conceptual puzzles of the theory of radiation and the constituents of matter, mentioned above; it also made possible further developments in the theory of matter, leading to our current standard model of particle physics, the story of the basic constituents of matter. One of the major features of the quantum mechanics is the principle of uncertainty, as first proposed by Werner Heisenberg, which argues, in part, that the exact position of any rapidly moving charged object like an electron in an atom can never be located with certainty—instead, the electron's position can be determined probabilistically. Hence, the thought arose that a typical electron in an isolated atom is like a cloud with very dense regions of the cloud representing probable places where the chance of finding an electron is most likely, and much less dense regions where this electron is least likely to be found. Another significant feature of the field of quantum mechanics is *discreteness*. For example, a specific electron in an isolated atom can assume only a few particular types of motions, called *states*, and particular values for its energy to stay in those states or levels which are likely, called *energy levels*. The glaring discrepancy in explaining black body radiation in the UV range of the spectrum also, known as the *ultraviolet catastrophe*, was rightly and clearly solved by the new specialized theory of quantum mechanics. Thus, it was during the first 30 years of the 20th century that the basic conception and evolution of quantum mechanics took place. In 1900, Max Planck introduced the basic ideas of quantum mechanics. He was thereafter awarded the Nobel Prize in physics in the year 1918 for discovering the quantified nature of energies, or *quanta*. The quantum theory was more prominently accepted after the Compton effect clearly showed that light indeed carries momentum and can thus scatter particles. During this time Louis de Broglie established his theory of wave-particle duality, which furthermore paved the way for the acceptance of the critical theories of quantum mechanics. In 1905, Albert Einstein successfully established his

theory on the photoelectric effect by using quantum mechanics and again in the year 1913, Niels Bohr effectively explained the stability of Rutherford's model by using quantum theories. Thus, the general quantized theory of the atom gave a new dimension for the beginning of an era of full-scale quantum mechanics around the year 1920. Heisenberg, Max Born, Paul Dirac, Wolfgang Pauli, and Erwin Schrödinger were some of the eminent physicists who extensively participated in the development of quantum mechanics during the 20th century. Heisenberg's uncertainty principle, proposed in 1925, asserts the impossibility of simultaneously and precisely measuring the momentum *and* the position of a moving charged particle; that is, it's possible to measure one or the other, but not both. Niels Bohr's ("Copenhagen") interpretation of quantum mechanics denied the possibility of any type of fundamental causality in quantum mechanics. In its new form, quantum mechanics became an almost indispensable tool in the explanation and investigation of the different phenomena at the atomic level. In 1920, Sir Satyendra Nath Bose's work on the *photon*—a particle representing a quantum of light—and its significant role in quantum mechanics led to the foundation of Bose-Einstein statistics, the theory that describes the fundamentals of the Bose-Einstein condensate. After this interesting discovery in quantum mechanics, the *spin-statistics theorem* emerged, which gave rise to the possibility of a particle of either being a *boson* (i.e., it follows Bose-Einstein statistics) or it can be a *fermion*, which follows the Fermi-Dirac statistics. Later on, it was found that all the fundamental bosons can actually transmit forces, like the photons, which are capable of transmitting light.

Subatomic Forces and Particle Physics

In the early 18th century, three fundamental forces were known to science: gravity, the electric force, and magnetism. By the very end of the 19th century, however, views had changed, and now it was understood that there were actually only two types of fundamental forces: the electric and magnetic forces are thoroughly unified into one fundamental force, which is called the *electromagnetic force*; this can be counted as one of the greatest achievements of 19th-century physicists. It turns

out that all magnetic fields are created by the rapid motion of charged particles, and these charges are perhaps the basic source of the electric force. Hence, in 1899, many scientists thought that there exist only two major fundamental forces: the electromagnetic force and gravity. And since these physicists had a fairly thorough understanding of both of these forces; they thought that they knew all of it. But the story took a new turn, specifically during the mid-20th century, when two new types of fundamental interactions were discovered. They were found to be subatomic—meaning that they act at scales much smaller than an average atom—and to reside inside the nucleus. The strong nuclear force binds the neutron and the proton to form the nucleus. This force also holds together the respective protons and neutrons in a nucleus. Similarly, a weak sub-nuclear force is responsible for the certain radioactive decay of nuclei as observed in case of many elements. In 1969, the three pioneer physicists Sheldon Glashow, Abdus Salam, and Steven Weinberg succeeded in a process of unifying this weak forces with electromagnetism. Thus, the basic fundamental forces are the strong nuclear force, electromagnetic interactions, and the force of gravity.

Special Relativity

Newton's laws of classical mechanics largely fail to explain phenomena at short distances and at high speeds, which can be correctly explained by quantum mechanics. At the very beginning of the 20th century, Albert Einstein developed the special theory of relativity, which greatly explains the dynamics of objects traveling at extremely high speeds. The effects of the laws of special relativity are only applicable for things moving at a velocity approaching the speed of light. One typical effect that has come to be widely known is the mass–energy equivalence, as embodied by the famous mathematical equation $E = mc^2$, meaning that energy equals mass times the speed of light squared. An enormous amount of energy can therefore be produced by the destruction of a small amount of mass. This not only explains the vast amount of energy stored in stars but also served as the basis of atomic bombs. One of the very interesting consequences of this special theory of relativity is the unification of space and

time into a typical four-dimensional world. Of the three papers published by Einstein in 1905, the third, titled “On the Electrodynamics of Moving Bodies,” clearly and strongly laid out his theory of relativity, which would change the course of physics. The theory was actually derived from Einstein's perspective on the luminiferous aether. Einstein realized that if the aether actually, exists it is more of a medium in which electromagnetic waves travel. The aether also provides a fixed background or a perfect reference frame against which all the motion that exists in the universe can be effectively measured. Any arbitrary point in the universe can be taken to be as the origin where three perpendicular axes, designated as x , y , z , meet. Thus, any particular point in the universe can be defined by three numbers, thereby specifying its actual distance from the origin along with the three axes. To clearly describe the path of a moving particle, one would only need to know the exact values of all three numbers at different intervals of time. The aether was considered to be the absolute frame of reference, as it was thought to be stationary, while any object or wave would move with respect to this stationary frame of reference. Scientists on Earth can only measure the absolute relative motion of any object with respect to the instruments stationed on earth in an absolute reference frame. Thus, to determine the absolute motion of that particular moving object, these scientists would need to accurately measure the absolute motion of each of these instruments with respect to the aether. Initially, scientists believed that the earth was actually the unmoving center of everything, but they realized that earth was actually part of a larger solar system, and they placed the Sun as the unmoving center of the universe. With newer perspectives, the view of space and time changed and so did the mathematical interpretations governing Maxwell's equations and Newton's laws of motion. In a postscript to his 1905 paper, Einstein raised the question of whether the inertia of a body depends on its energy content—which can be considered the most dramatic surprise of the theory of relativity at that time. Electromagnetic energy must clearly travel at the constant speed of light, but anything which has mass can never achieve that much speed irrespective of the magnitude of force acting on it for whatever duration. Thus, the higher an object's

velocity is in the observer's reference frame, the more magnitude of force must be exclusively applied to increase its speed by the requisite amount. The work done on it in this case actually causes the mass or inertia of the object to increase considerably.

General Relativity

During the later years of the 19th century, physicists believed that the universe is actually made up of only thousands of stars, which are not arranged in any particular pattern. The most distant stars were actually thought to be about 100,000 light years away from the earth. However, during the 20th century, with the development of larger and much better telescopes, astronomers were able to observe objects that were about 10 billion light years away. Furthermore, they have found that the universe actually contains a large number of interesting structures. It is not until the year 1920 that scientists recognized the existence of galaxies, which are large collection of distinct stars grouped together in a relatively localized area of the universe of about 100,000 light years in size, usually containing one hundred billion stars. Some of these galaxies are pancake shaped, while others are spiral shaped and are referred to as spiral galaxies. Some galaxies are ellipsoidal, and many others are irregular. The galaxy in which the Earth and the Sun reside is called the Milky Way, named after the band of whitish hue that can be observed from Earth on a clear night. The next two large galaxies nearest to the Earth are the large and small Magellanic clouds, which are irregularly shaped and can be only observed from the Southern Hemisphere of the Earth. Andromeda is the third nearest galaxy from Earth; it is spiral shaped. During the last three decades astronomers have made a significant finding that, occasionally, 100 to 1,000 galaxies group together to form galaxy clusters. They also found that there are regions in the universe which are devoid of any galaxies; these regions are called *giant voids*. In trying to understand how to make his special theory of relativity fit with gravitation theory, Einstein developed the general theory of relativity, a theory that explained much of what we see in the large-scale structure of space, including the astronomer Edwin Hubble's observation

of the difference in the wavelengths of similar stars at different distances from the Earth and his postulation that the most distant stars are moving away from the Earth at greater speeds.

Einstein's second revolution in 20th-century physics is general relativity, his theory of gravitation. The general relativity theory provides deeper insights into the actual nature of gravity. Developing the general theory of relativity took Einstein to unusual mathematical territory. A simple experiment might provide a clear perspective of this theory and its analysis. Let us suppose that in a laboratory an observer is involved in measuring the motion of some falling bodies. There is no air resistance, because the bodies are in separate vacuum chambers within the laboratory, and the bodies are magnetically and electrically neutral. The observer's only goal is to measure the force of gravity on these bodies, which are free from all other forces. The observer learns that the velocity of these bodies gradually changes in a typical way, and it is the same for all such bodies regardless of any difference in their masses. The vertical motion of these bodies continuously increased in the downward direction at a rate of 9.8 meters per second, while their velocity parallel to the ground remained constant. The observations made during these experiments produced a very clear conclusion. The laboratory, along with the observer is, actually in a gravitational field with an acceleration of 9.8 meters per second, or 32 feet per second, simply called *g*. However, if an observer is considered to be on the falling body, the whole perspective would change. To an outside observer, the laboratory and the lab observer were viewed as accelerating upwards with a force of one *g*. Thus, the main fact is if neither of the observers was allowed to see outside into the surroundings, they would never be able to conclusively predict whether they are in any accelerated frame of reference or in a gravitational field. Pursuing this thought further, Einstein was able to unify space and time together into a four-dimensional space-time grid. People usually envisioned position as defined by a three-dimensional geometry or space, as a grid of imaginary meter sticks which basically stretch themselves to infinity in three different directions perpendicular to each other in space. But Einstein argued the presence of a fourth dimension: *time*,

denoted by the t axis through which everything gradually moves at a specific rate of one meter per second. Now, to distinctively envision Einstein's concept of space-time, a four dimensional grid with markings x , y , z , and t can be imagined. Photons have no specific mass or else they would have to travel much slower than the speed of light according to the concepts of the special theory of relativity. When Einstein published his general theory of relativity in the year 1915 many physicists were seriously determined to test his predictions. Fortunately, the solar system, on rare occasions such as a solar eclipse, provided a way to prove these predictions true. When starlight on the way to Earth passes close to the Sun, solar gravity actually diverts the path of the light by a measurable amount. However, during a total solar eclipse, astronomers are capable of viewing and measuring the distinct pattern of stars that are normally invisible in a bright sky. The eclipses rarely occur—only once or twice a year—but they usually occur along very narrow paths as the shadow of the Moon sweeps across the surface of the Earth. In 1917, when Einstein was delicately exploring the implications of his newly described mathematical theory of relativity, he discovered that his theory had predicted the existence of a continuously contracting or expanding universe. At this time he also noticed that the exquisite solution to his mathematical equations actually included a typical and apparently arbitrary value termed *the cosmological constant*, which can ascertain different values leading to different rates of contraction or expansion. Only one particular value can lead to stability, and that was the value on which nature had settled. Although Einstein considered the presence of the cosmological constant as his greatest mistake, the term eventually made a comeback at the very end of the 20th century and is still deeply intriguing physicists. Now physics has created more problems than it has actually solved, starting from Einstein's theories of relativity to Neils Bohr's structure of atoms; new areas have to be opened, leading to a full-scale effort to virtually reestablish physics on totally new fundamental principles. Einstein's general theory of relativity, proposed in the year 1910, concluded that space is actually highly curved and finite in size. Einstein, by expanding relativity to the accelerating

reference frame, posited an equivalence between the force of gravity and the inertial force of acceleration.

The History of the Universe

Toward the end of the 19th century, scientists thought the age of our universe was about several hundred million years. Moreover, very little was known about the history of the universe. Today it is understood that universe began as an almost inconceivably hot concentration of energy and mass. As time advanced, the concentrated universe expanded immensely, which means that the basic fabric of space was highly stretched, and through this stretching, the material was rapidly dispersed and the universe eventually cooled. Finally, gravity affected very high concentrations of matter, enabling them to collapse rapidly into individual galaxy clusters at very large scales and others into stars at much smaller scales. This process of star formation with the aid of gravitational collapse also continues today, typically at a much slower rate. The application of terrestrial physics to the rest of the universe, to be able to infer its time evolution and to understand the mechanisms of that evolution marks a significant moment in the development, expansion, and application of theory.

Other Significant Advances in Theory

In addition to the revolutions of its early part, the 20th century was marked by significant advances in the nature of the theories it employed.

The theories in this domain of *nonlinear and chaotic systems theory* are important for their practical results in understanding, for example, weather, climate, fluid flow, and scaling phenomena. They are also important because they mark a significant change in our understanding of the relation between determinism and prediction, which while conceptually distinct had always been connected in most discussions of physical phenomena. Chaotic systems are among the non-linear types. They have a peculiar feature that while their equations of motion, and the solution spaces of those equations, are completely deterministic, they are highly resistant to prediction and indeed it can be shown that in many regimes their exact

behaviors are *in principle* impossible to predict. The slogan here is that no matter how small a change one makes in initial conditions, there can be arbitrarily large differences in final outcomes. The significance of this is that, since our observations are, in principle, of finite precision, then no matter how precisely we can determine the initial conditions, the outcomes will be unknowable. This kind of unpredictability is totally different from that introduced by quantum theory, where the suggestion in most quarters is that systems it studies are indeterministic. But chaotic systems—fluid flows for example—can be totally determined and yet unpredictable even so. (For more on this, see the entry on Complex Systems.)

A second major theoretical development is in the connection between thermodynamics and the microstructure of matter studied in kinetic theory. This connection and theoretical advances that resulted from studying it had important results in thermodynamics itself as well as profound effects on theory in solid-state physics, on gravitational theory, as well as on perturbation theory.

The Twentieth Century saw important advances in physics that shaped and continue to shape our understanding of the world, as well as our understanding of physical theory and its application to the world.

Syed Feroj Ahmed

See also Complex Systems; Cosmology; Physical Theory; Physics, 19th century; Physics, Contemporary; Physics, Gravitation Theory; Physics, Quantum Theory; Physics, Thermodynamics

Further Readings

- Blum, A. S., & Rickles, D., (Eds.). (2018). *Quantum gravity in the first half of the twentieth century: A sourcebook*. Edition Open Access, Max Planck Institute for the History of Science.
- Crease, R. P., & Goldhaber, A. S. (2014). *The quantum moment*. W. W. Norton.
- Einstein, A. (2005). *The meaning of relativity*. (5th ed.) Princeton University Press. (Originally published 1954)
- Einstein, A., & Infeld, L. (1938). *The evolution of physics: The growth of ideas from early concepts to relativity and quanta*. Cambridge University Press.
- Fermi, E. (2000). *Thermodynamics*. Dover.
- Holland, P. (1995). *The quantum theory of motion*. Cambridge University Press.
- Mehra, J., & Rechenberg, H. (1982). *The historical development of quantum theory*. Springer-Verlag.
- Munitz, M. K. (1986). *Cosmic understanding: Philosophy and science of the universe*. Princeton University Press.
- Planck, M. (1997). *Eight lectures on theoretical physics*. (A. P. Wills, trans.) Dover Books. (Originally published 1909)
- Rae, A. (2012). *Quantum physics: Illusion or reality?* Cambridge University Press.

PLATE TECTONICS

Plate tectonics is not only the central theory of earth science but also an extraordinary case for those wishing to understand the social acceptance of a theory, as its main postulates circulated for a long time but did not generate consensus. Plate tectonics theory considers that the Earth is composed of layers, and the rigid outer layer, known as the *lithosphere*, is broken up into several plates that are in movement. The theory emerged in the 1960s anchored in two principles: seafloor spreading and continental drift. It explains how mountain ranges, earthquakes, volcanoes, and continents fracture, how oceans form shorelines, and other features. The impact of the theory of plate tectonics to geoscience has been compared to that of Darwin's theory of evolution on biological sciences. It is said to have deployed a new paradigm that profoundly changed the dominant ideas about how the Earth works.

The geophysicist Alfred Wegener is usually recognized as the precursor of the theory, for his development of the continental drift theory in 1912. But it took 50 years until his ideas about the movement of the continents were vindicated. As will be described and discussed in what follows, plate tectonics theory is not only a major geoscience theory, but its history is also a controversial case in which rationalist and constructivist approaches argue about why it was so difficult for scientists to accept that the landmasses were actually moving.

The Emergence of Plate Tectonics Theory

The understanding of the structure of the seafloor was one of the pillars of the theory development and, perhaps unexpectedly, war became an ally for its purpose. German submarines caused heavy losses to the allied forces in World War II, which prompted the U.S. Navy to improve knowledge in geophysics and oceanography to provide means to detect enemy submarines. Many important scientific discoveries took place within this context. The U.S. Navy asked Maurice Ewing, then working at the Woods Hole Oceanographic Institution, to investigate why sonar equipment tended not to work well in the afternoon. This led Ewing to realize that temperature effects were bending the sound waves in such a way as to create a *shadow zone*, where sonar transmissions went undetected. This knowledge was then used by the Navy to better detect enemy submarines. Harry Hammond Hess, an American geologist who would have a major role in the development of plate tectonics theory, was also involved in war affairs. He was captain of a ship in charge of echo-sounding of the Pacific basin. Hess discovered that bottom of the Pacific floor was not smooth as expected, but full of hundreds of flat-topped mountains. Intriguingly, these mountains looked like their tops had been eroded by wind, but they were actually a few kilometers under water. Hess proposed an audacious hypothesis: Those mountains under the ocean were drowned ancient islands.

The following years saw a reinforced encouragement to geophysics, particularly within the context of the cold war. The leaders of the United States, of the Soviet Union, and of the United Kingdom reached an agreement to ban further testing of nuclear weapons. The problem was the undeveloped capacity to test underground violations through seismology. The solution came into the shape of a new organization that would improve earth science: the World-Wide Standardized Seismographic Network. Funded by the U.S. government, this global network deployed approximately 120 seismographic stations that generated an unprecedented amount of seismic data. At the same time, the Soviet Union deployed a similar system through the Unified System of Seismic Stations, with its own 168 seismic stations.

Other scientists, such as Edward Bullard and Patrick M. S. Blackett, also got involved in war affairs, mainly with geomagnetism research. All of them were aware of the continental drift theory and had it in mind while analyzing their data. By the mid-1950s, Blackett and his group in London argued that the variations in the earth's magnetic field showed that rocks had not remained stationary relative to Earth's magnetic field: This variation would support the theory of continental drift. Earth's ancient magnetic field can be traced in rocks because minerals indicate, like any magnet suspended in Earth's field, the direction to the magnetic pole at the time the minerals were crystallized or deposited. Studies in the 1950s showed that each continent yielded different results, but these differences could be reconciled into the same path by envisioning the joining the continents as suggested by Wegener. In other words, this paleomagnetic analysis implied that the poles' geographic variations could be explained by the wandering of the continents. A few years later, Hess supported Blackett's point of view by adding an explanation: Mantle convection currents might drive continental drift.

For centuries, it has been known that there are different kind of rocks in the Earth's interior. The Earth is made up of three main components, or layers, which vary by composition and behavior: the crust, the mantle, and the core. A relatively thin crust sits on top of the mantle. The mantle is much thicker than the crust; it contains 83% of Earth's volume and has a depth of 2,900 km (1,800 miles). Finally, beneath the mantle is the core, which extends to the center of Earth, some 6,370 km (nearly 4,000 miles) below the surface. There are two types of crust, *continental* and *oceanic*. The outer shell of our planet (crust and upper mantle) consists of the coolest part of the Earth, called the *lithosphere*. According to plate tectonics theory, here is where the plates are: The lithosphere is broken into several separate, rigid blocks, or plates. Below is the *asthenosphere*, a much hotter and more fluid layer. Here is where convection currents are generated; because of its ductility, the asthenosphere is able to deform or stretch. Tectonic plates rely on this viscous layer, which allows them to move. Convection currents generated in the asthenosphere push magma—molten or semi-molten rock—upward through volcanic vents and

spreading centers to create new crust. And by putting stress on the lithosphere above, they may also provoke earthquakes. According to Hess, mantle convection currents are the force that moves the plates. Where the edges of the plates meet, earthquakes take place. But several other things may happen as well. If those plates carry continents, which are lighter than the ocean floor, they may collide, causing high mountains to rise.

Hess postulated that molten rock from Earth's mantle emerges from active volcanoes in the ocean. As the magma cools, it spreads, creating a successively younger ocean floor and pushes apart the continents, as new ocean crust spreads away from the ridges. Hence, the ocean floor is younger than the continents. This process is called *seafloor spreading*, and it creates new crust. Samples collected from the ocean floor indicate that the age of oceanic crust increases proportionally with distance from the spreading center. These age data also allow researchers to determine the rate of seafloor spreading, which varies from about 0.1 cm (0.04 inches) per year to 17 cm (6.7 inches) per year. So new ocean floor is continuously being created. But tectonic plates may also crash into each other at *subduction zones*, where the edge of the denser plate slides beneath the less dense one, melting back into the Earth's mantle. Subduction destroys old crust. In a sort of balance between spreading and subduction, the shape and diameter of the Earth remains constant. All these processes with the oceanic crust are accompanied by much volcanic activity and by earthquakes as the crust repeatedly rifts, heals, and rifts again.

Hess's report, "The History of Ocean Basins," was formally published in 1962. Overall, Hess supported Wegener's theory of continental drift, as Hess reasserted that the once-joined continents had separated into the seven that exist today, and they had done so because they are transported by the shifting tectonic plates on which they rest.

The Continental Drift Controversy

Through plate tectonics theory we think about the Earth as a dynamic structure, with continents moving away from each other. For centuries the Earth was conceived as a rather stable structure, with no internal movements at all, or at most with

vertical motions (volcanoes and sinkings), but sideways movements of Earth's layers was not an option. Plate tectonics has transformed the way we view the surface of the Earth. But this revolutionary conception of the Earth was already present or implied at the time of continental drift theory. Proponents of plate tectonics, including Hess, recognized Alfred Wegener's continental drift theory as the forerunner of their work.

Born in Berlin in 1880, Wegener obtained a doctorate in astronomy in 1905. He then pursued a strong interest in the field of meteorology and polar research. In 1906, Wegener participated in the first of his four Greenland expeditions, where he took meteorological measurements in the Arctic climatic zone. He then obtained a position at the University of Marburg, lecturing in meteorology and astronomy. He first became aware of the problem of the continents' structure and was inspired to propose the initial sketches of what would later become his continental drift theory as a result the impact of his having noticed the complementarity of the Atlantic coastlines on a map in 1910. From that moment, he carried out a series of readings that influenced his ideas even more. Previously, in his book *The Face of Earth* (1883), Eduard Suess, an Austrian geologist, had brought to his colleagues' attention the very large dislocations of the Earth's crust. Suess even coined the word *Gondwanaland* for a proposed paleo-supercontinent consisting of Africa, South America, India, and Australia. Finally, in 1912 Alfred Wegener published his theory of continental drift, 50 years before Harry Hess was to reaffirm it with the plate tectonics theory.

Wegener argued that several millions of years ago, all continents were united in one supercontinent that he called *Pangea*. Later on and slowly, during the Cretaceous period, Pangea began to break apart. At that period, what we now recognize as the American continent began to drift away from Europe and Africa, while Asia separated from the Antarctica and South Africa by rotating counterclockwise. The movement didn't stop, as continents continued to drift away from one another, causing, for instance, the separation of Australia from Antarctica. In 1915 Wegener published his theory in a more developed way, in his book *The Origin of Continents and Oceans*. The title itself was quite explicit about the

explanatory power of the horizontal displacement of the continental masses.

Continental drift theory wasn't just a bold speculation. Wegener provided a diversity of evidence supporting his theory. He argued that the continents' margins matched, suggesting that they had once been united and then split apart. He also analyzed the distribution of similar fossils and current flora and fauna, which suggested extensive interchange of fossil species between continents during certain geological periods (in other words, what is today a genus or family was once a species in prehistoric times, therefore there must have been exchanges that could only have taken place if the lands had been connected). The several geological structural similarities among the continents as well as the coastline location of mountain ranges were additional elements that supported his theory. His own background and experience also proved useful to bolster his theory, as the longitudinal measurements from the Danish Greenland expeditions indicated a present-day westward displacement of Greenland; and geological indicators of paleoclimatic change were another key element, as he argued that the observed drastic changes of the climates linked to the shift of the continents relative to the poles. In short, Wegener marshaled evidence from very different domains to support his theory of continental drift.

Still, his theory wasn't accepted in his lifetime—not because it had gone unnoticed, rather the other way around: It got the attention of his colleagues, who mostly rejected it. Why was continental drift theory rejected by the scientific community in 1912 to be accepted only 50 years later? There are two general and different ways of explaining this phenomenon: rationalist and constructivist perspectives, both of which have provided frameworks for understanding this controversy.

The Controversy From a Rationalist Perspective

When trying to explain the controversy around the continental drift theory, many philosophers argue that although Wegener was right, the scientific community had legitimate reasons not to accept his theory. Some stated that Wegener had the general idea right, but he lacked enough

evidence to convince his colleagues. The main flaw would be an explanation of how the continents displace horizontally. This sort of evidence would come only 50 years later, with the discoveries that allowed Harry Hess to postulate the seafloor spreading hypothesis, which would serve as an explanation of the continental drift. This implies that Alfred Wegener had a privileged intuition that allowed him to recognize what was to become a correct theory but without the sufficient evidence to support it. This sort of characterization is summarized in the expression “being ahead of his time.” In that sense, Wegener is vindicated by these perspectives as an early advocate of an immature theory, “born at the wrong time,” or facing fierce opposition because he made a discovery before its correctness could be confirmed. At the same time, this explanation of the controversy covers in rationality the rejection adopted by Wegener's colleagues, as it would be a rational response not to accept a theory that lacks enough evidence.

Those points of view tend to explain the social acceptance of a theory by arguing that a theory gets accepted because it is true or because it is supported by sufficient evidence. But there are also rationalists who have perceived that many theories accepted in the past can today be considered incorrect, so the notion of truth maybe more unstable than is commonly supposed. Those rationalists may argue that theories are accepted because they offer more precise explanations than other available ones, or because they more completely solve those problems than others. Other who take this perspective believe that a theory can be adopted for its ability to capture the facts it chose to explain, its degree of generality, or its ability to focus on the things we are particularly interested in understanding. Ultimately, rationalist perspectives aim to explain the social acceptance of a theory by seeking an epistemic or cognitive element that supposedly holds the intrinsic superiority of that theory and induces the scientific community to accept it. The continental drift debate has been a fertile ground for those explanations, as many tried to explain the opposition that Wegener faced when he published his continental drift theory by arguing that it didn't provide enough evidence to be convincing. Others, within a similar perspective, explained the social rejection of the theory in terms of some

internal flaws or incompleteness, like the lack of a convincing explanation for the displacement of the continents. All of them consider that the scientific community rejected the theory in the early years of the 20th century because they had epistemic reasons to do so, and that the revised version of the theory under plate tectonics provided better epistemic elements (empirical evidence or more solid explanations) that forced the scientific community eventually to accept it.

The Controversy Through Constructivist Perspectives

A different approach to the acceptance of the continental drift theory has also been deployed, one more focused on the cultural and social contexts that shaped this debate.

One of the first implications that arise when looking at the cultural context of the debate is that the ideas about the displacement of the continents didn't start with Wegener in 1912, but long before. An early record of an advocate of the continental drift can be found in Abraham Ortelius, a 16th-century geographer, who included two central aspects of the continental drift theory: the continental fit and the movement of land masses. In his 1596 book, titled *Thesaurus Geographicus*, Ortelius argued that America had been torn away from Europe and Africa, and that the vestiges of the rupture can be seen from the complementarity of their coasts. The list of ancient proponents of the theory includes notable names such as Alexander von Humboldt, Francis Bacon, Comte de Buffon, François Placet, Thomas Young, Richard Owen, Jean-Baptiste Lamarck, and Antonio Snider-Pellegrini. All of them, among others, have proposed, with different degrees of depth, that the continents as they exist today once comprised a single continent, as can be inferred by the shape of their coasts as well as by such evidence as the similarities of the fossils found in the various continents. More closely contemporary authors to Wegener (but still before him) also developed continental drift theories, such as Frank Bursley Taylor, Roberto Mantovani, and William Pickering. The case of Taylor is revealing: By the end of 1908, this American geologist argued that, in the Tertiary period, the continents had moved

because of what he called *deforming forces*; 2 years later Taylor provided geological evidence in support of his theory. In earlier years, in fact, it was common to read about the *Taylor–Wegener hypothesis*, although later the name of Taylor was erased from the origins of the continental drift theory, jointly with the rest before him. Although Wegener provided an elaborated and developed notion of continental drift, the assumption that he should be considered the forefather of continental drift theory relies on the way geology as discipline constructed its own past in the 1960s, not on historical records of the pioneers of the theory. In any case, the idea that continents were not something fixed on Earth, but moving horizontally, was a persistent and recurring idea for centuries. These mobilist thinkers and scientists were opposing a dominant point of view, that of the fixists.

If mobilists conceived things to be in permanent change, fixists tended to believe that things had remained in the same place since the Earth's creation. At least until the middle of the 20th century, most biogeographers were fixists, as they considered that the positions of the continents had been relatively stable over the time scales. The extent, depth, and development of mobilists' theory could be very diverse, but there was always at least the idea of the displacement of the continents and some kind of evidence to support it. And they were always rejected by a majority of fixists. To reject is not simply to ignore, so they had to provide some answer to the arguments and evidence brought by mobilists. And they did. For instance, fixists managed to explain the similarities of the fossils that existed in different continents, arguing that the continents were fixed always to the same place, but also were once connected by trans-oceanic land bridges. While mobilists explained the existence of mountain ranges as the result of the crush of continents, fixists considered that the Earth was cooling and contracting from a previously molten state, like a baked apple coming out of the oven. For fixists, this contraction would explain the great mountain ranges all around the world. In every case, for fixists, it was impossible to conceive that the continents could move horizontally, and they denied the evidence and reasoning put forward by mobilists.

To view the debate around the continental drift theory as part of a broader cultural background that shaped it implies the recognition that these two styles of thought—mobilism and fixism—had coexisted for centuries, providing a general framework for the acceptance or rejection of the continental drift theory. There were many people before Wegener who had entertained similar ideas. Thus, Hess's ideas had their predecessors; for instance, he incorporated the idea proposed by Swiss geologist Emile Argand in the 1920s that mountain belts are created when two continents collide. Such ideas had a few supporters but also a lot of detractors. Although Wegener himself had some supporters, most of his colleagues rejected his theory. So what changed in the 1970s when the idea of continental drift was widely accepted? According to rationalists, some new evidence or perfected theory turned it irrefutable. But when Hess and his contemporary colleagues presented the plate tectonic theory, the scientific community didn't simply adopt it. By the middle of the 1960s, some prestigious scientists, such as Gordon MacDonald, J. Lamar Worzel, and Sir Harold Jeffreys, opposed the theory, disputed the new evidence, and labelled the plate tectonics theory as biased propaganda in favor of continental drift. In 1972, Paul S. Wesson collected 74 objections to continental drift and plate tectonics and concluded that the continents have never moved with respect to each other. Moreover, a survey of geologists, conducted in the 1960s and 1970s, revealed that those who already were convinced of the theory of continental drift viewed Hess's findings favorably, while those who were against the continental-drift theory rejected those same findings.

That the theory was eventually accepted may be attributable largely to social elements. Most likely, new generations of geologists, educated under the influence of mobilists who had attained positions of authority at the universities, accepted the theory just by inheriting the mobilist style of thought. Indeed, those who accepted late tectonic theory during the 1960s and 1970s were mainly young scientists. So in the end, most scientists did not surrender to some irrefutable evidence or reasoning; nobody changed their mind, but rather the change was due to a shift in social influences: mobilists increased the number of young followers.

The Importance of Plate Tectonics Theory

Plate tectonics theory is considered a fundamental scientific theory, capable of explaining most of the structural phenomena on Earth, such as earthquakes and the origin of continents. The idea of continental drift has acquired consensus among the scientific community, and it is used to explain not only the history of Earth landmasses but also the planet's evolution. The continents bordering the Atlantic Ocean, for example, are believed to be moving away at a rate of 1 or 2 cm (0.4–0.8 inches) per year. The European and North American plates are moving apart very slowly, but millions of years in the future, parts of California and Mexico will probably drift off to become an island. The emergence of new mountain ranges, islands, and associated earthquakes are possible to predict or at least to think, due to plate tectonics theory.

That Wegener had proposed a continental drift theory in 1912 without receiving any support from the scientific community has led to comparisons between the Wegener/plate tectonics situation and Gregor Mendel's being credited for discovery of genes. Published in 1866, Mendel's research was largely ignored until 1900, when he began to be seen as the father of genetics. This significant interval between the first presentation of the theory and its later acceptance is often explained as Mendel's being far ahead of his contemporaries' thinking. Similarly, many philosophers of science tend to explain that Wegener was ahead of his time, and that plate tectonic theory, proposed 50 years later, was more convincing. This sort of whig history (a historiography of science that assumes that rejected or ignored theories were just not as good as present-day theories, in a sort of linear progression of science) has placed all debates surrounding the acceptance of the theory as mere rational and epistemic issues. In a much subtler analysis of Mendel's laws of heredity, Augustine Brannigan considers that Mendel was not the founder of genetics nor was his contribution overlooked; he argues that the contexts in which Mendel's work was presented as well as its significance changed substantially over time. In a similar way of thinking, the

cultural background surrounding the long-term debate about the movement of the continents is what may have also changed, turning it finally into an acceptable idea, thereby enabling plate tectonics to develop.

In any case, plate tectonic theory is not only a fundamental scientific theory that explains many issues in earth sciences, but also an important historical case that enables us to understand what leads a scientific community to accept a theory, and when an idea comes to be seen as socially acceptable.

Pablo A. Pellegrini

See also Constructive Empiricism; Epistemology; Evidence; Philosophy of Science; Social Construction of Scientific Knowledge; Theory Change

Further Readings

- Frankel, H. R. (1979). The reception and acceptance of continental drift theory as a rational episode in the history of science. In S. H. Mauskopf (Ed.), *The reception of unconventional science: AAAS selected symposium* (pp. 51–89). Westview Press.
- Frankel, H. R. (2012). *The continental drift controversy. Volume IV: Evolution into plate tectonics*. Cambridge University Press.
- Frisch, W., Meschede, M., & Blakey, R. (2011). *Plate tectonics: Continental drift and mountain building*. Springer.
- Greene, M. T. (2015). *Alfred Wegener: Science, exploration and the theory of continental drift*. Johns Hopkins University Press.
- Hess, H. H. (1962, Nov. 1). History of ocean basins. In A. E. J. Engel, H. L. James, & B. F. Leonard (Eds.), *Petrologic studies* (pp. 599–620). Geological Society of America. <https://doi.org/10.1130/Petrologic.1962.599>
- Messeri, P. (1988). Age differences in the reception of new scientific theories: The case of plate tectonics theory. *Social Studies of Science*, 18(1), 91–112.
- Oreskes, N. (Ed.). (2018). *Plate tectonics: An insider's history of the modern theory of the earth*. CRC Press, Taylor & Francis Group.
- Pellegrini, P. A. (2019). Styles of thought on the continental drift debate. *Journal for General Philosophy of Science*, 50, 85–102. <https://doi.org/10.1007/s10838-018-9439-7>

PRAGMATISM

Pragmatism, a philosophical movement often portrayed as U.S. America's leading contribution to the history of Western thought, originated in and around Harvard University during the late 19th century. The classical pragmatists include its founder Charles Sanders Peirce, its popularizer William James, and its key 20th-century representative John Dewey. Its main theoretical contribution is found with its use of the pragmatic *maxim*, a method or principle that emphasizes the clarification of theories, concepts, and hypotheses by establishing their connections to human practice and experience. With their discussion of the importance of practice, the pragmatists might be seen as largely hostile to theory. However, in both its classical and contemporary forms, pragmatism continues to contribute to theory in philosophy, education, sociology, and political theory. Although frequently critical of theory and its uses in philosophy, pragmatists generally seek an improved understanding of the relationship between theory and practice. In the process, they argue that the relevance of theory in its various formulations must be measured in terms of its practical consequences.

It was Peirce who first introduced the pragmatic maxim. James would notoriously adapt this maxim to his own purposes, leading to its most controversial formulation with the pragmatist theory of truth. Dewey's later experimentalism uses key pragmatist insights to argue for a rejection of any strict distinction between theory and practice. Among neopragmatists, Richard Rorty is the most famous for his general aversion toward the use of theory in philosophy. Recent work in political theory has looked to both Dewey and Peirce in the attempt to provide a defense of democracy that is grounded in pragmatist epistemology. As will be discussed in the following sections, these contributions are unified in sharing an epistemological perspective that is *fallibilist*, recognizing that any single belief could be in error, while looking to also clarify the norms that govern all forms of inquiry, stretching from the scientific to the moral and political.

The Pragmatic Maxim and Theory of Truth

Pragmatism emerged out of discussions between Peirce, James, and other like-minded intellectuals at what they called the *metaphysical club* in Harvard around 1870. Peirce further developed his contribution to these discussions in his publications during that period, but it was not until the turn of the century that James popularized the term *pragmatism* in public lectures. Both Peirce and James think of pragmatism as involving a method, principle, or maxim used to clarify concepts and hypotheses by linking them to their practical consequences. But they importantly differ in their understanding of the maxim and concerning how it should be applied in philosophy. Peirce insisted that his maxim be seen as a central logical technique required for progress within experimental science. Here, a statement is clarified when we list the experiential consequences we expect to occur if that claim is true. If one states that this diamond is *hard*, then this means that it will not be scratched by other substances. The concept *hard* has been clarified by indicating what practical effects we should expect to follow if that statement is true. This pragmatic degree of clarity is important for Peirce because of the central role in plays in inquiry. If our beliefs did not have any consequences, if their being true or false did not make any difference in our expectations, then the belief is empty from the standpoint of scientific inquiry.

James presents pragmatism as a method that helps to locate the significance of competing philosophical viewpoints and as a tool for highlighting empty verbal disputes. In his book, *Pragmatism*, he claims that the basic purpose of philosophy is to determine what difference the acceptance of one position would have at some specific point in our experience if this position were in fact true. Here, pragmatism is offered as a proposal, equipped with a method, which is justified through its fruitful application within the possible range of human experience. For example, he claims that the concept of freedom, once submitted to pragmatic standards, proves significant not through making actions morally accountable but because it allows for novelties in the world. The pragmatic meaning of freedom is located in

its promise for the possibility of improvement, something that we all count as a significant dimension of human life and experience. James is perhaps most notorious for his controversial remarks on pragmatism as a theory of truth. Here, he offers various formulations: the *true* as what is good for belief, truth as expedient for behavior, truth as what ideas become when they connect with other experiences and ideas as true instruments. These claims are open to both good and bad interpretations. They have been read by many critics as offering the untenable view that identifies truth with usefulness or the further surprising suggestion that a belief can be made true if it contributes to our happiness.

John Dewey's Experimentalism

In *The Quest for Certainty*, Dewey uses pragmatist ideas concerning the function of knowledge to criticize the strict separation between theoretical knowledge and practical judgment. He offers a historical diagnosis of the roots of this distinction, locating it in the Western tradition's philosophical pursuit of certainty. In response to the need for security in a perilous world, Greek thought held up the ideal of certainty as the ultimate measure of human knowledge. This view required that the objects of knowledge be fixed and unchanging and therefore something known independently of the changing environment that characterizes the arena of human practical activity. Dewey argues that this division between knowledge and action has been further preserved in modern thought, resulting in the view that beliefs about practical action are of less value than those achieved through intellectual theoretical pursuits. This separation between theory and practice has further resulted in diverting attention from addressing the practical problems of modern life.

Dewey further argues that this separation has been thoroughly rejected within the experimental perspective found in the procedures of modern scientific inquiry. He explores the profound changes in our conception of the theory–practice distinction that follows from the modern practice of knowing, illustrated within experimental science. Practical action should no longer be seen as inferior to theoretical knowledge, and the

commitment to security that is achieved through active control of our surroundings should be viewed as more important than any certainty found in theory. This does not result in the elevation of practice or action over theory. Rather, the important emphasis is on a deepened appreciation for the value of constant interaction between theory and practice. Dewey's pragmatic experimentalism takes both knowledge and practice as means for achieving goods and making them more secure within human experience. This further sets the main task for philosophy: Its function is to help encourage fruitful interaction between those beliefs arrived at through the most reliable methods of inquiry and practical concerns about those values, ends, and purposes that should direct human action and social life.

Rorty's Pragmatism as Antitheory

Rorty's controversial form of neopragmatism emerges from his critical analysis of the defining features of modern philosophy, where he focuses on dismantling the idea that knowledge involves the accurate representation of a world external to the mind. Taking inspiration from a nonstandard reading of Dewey, Rorty's pragmatism rejects its status as a theory of truth or meaning. Instead, he views pragmatism as denying the need for any epistemological or metaphysical view that would provide independent support for our human practices of deliberation and justification. For him, it more closely represents a critical attitude toward philosophical approaches that seek to justify claims in some type of grand universal philosophical theory. This does not amount to a rejection of the need for justification but of the view that our justificatory practices require additional validation from some independent philosophical source beyond them, as traditionally found in nature, God, or the mind.

This form of pragmatism, in its aversion to philosophical theory, further informs Rorty's political thought. Rather than seek philosophical principles that support democracy, he argues that our commitment to democracy has no need for such extra-philosophical support. This is based on his further controversial claim that the search for such support is incoherent, being based on an obsolete philosophical vocabulary. There are, he argues, no

external facts about human nature, reason, or morality that supply the needed foundation for democratic practices. Any defense of liberal democracy must then proceed from an acceptance of our historically contingent human practices of association and debate as the sole source of justification for those practices. Rorty's own defense of democratic ideals of freedom and justice will then hinge on descriptions of those practices that reveal their ongoing efficacy in addressing the practical political problems of our time and place.

Pragmatism and Democracy

John Dewey's view of the function of philosophy as addressing the conflict between scientific knowledge and human value resulted in the extension of pragmatist ideas into political philosophy and democratic theory. Recently, pragmatists have adapted Deweyan ideas in establishing an epistemic defense of democracy and democratic institutions. Others have been critical of this approach and instead look to Peirce for a pragmatist conception of epistemology that is capable of supporting a commitment to deliberative democracy.

This epistemic defense of democracy, whether in Deweyan or Peircean form, contains three main elements. First, it stresses the importance for democratic decision making of processes that ensure open public debate where citizens can offer information, distinctive viewpoints, and arguments. Second, it offers an epistemic defense of this decision-making process. Collective decision making is here presented as improving the results of this discussion; and acceptance of its outcomes is, at least, partially explained by its value for knowledge acquisition and problem-solving ability. Finally, this view gives a distinctly pragmatist account of the view of knowledge that underlies this deliberative process.

The neo-Peircean version of this view looks at the function the concept of *truth* places in our cognitive lives. Having a belief is to believe that a specific claim is true and to further accept that it would remain undefeated by the future course of experience and human inquiry. Being responsive to evidence, experience, and reasons is here taken as a *constitutive* norm of belief, meaning that it is part of the very nature of a belief that it be

sensitive to reason. This norm further points to the social nature of inquiry. Aiming at the truth requires that we test our beliefs against the available evidence where this involves the social practice of articulating and defending one's views to others. Seeking truth then demands that believers participate in the community of inquiry. Reflection on the nature of belief leads to the need for further inquiry involving community wide participation.

A commitment to deliberative democracy then emerges out of basic epistemic presuppositions of Peirce's pragmatist view of knowledge and belief. Having beliefs commits one to interpersonal norms of communication and discussion, where these further require recognizable democratic values such as freedom of speech, open debate, diversity of perspective, accountability, and openness to revision and change. If this type of inquiry is to function effectively, a further set of democratic institutions are needed to support forums for such deliberations, including the necessary feedback mechanisms in the form of periodic elections and a free press. It remains an open question whether this Peircean perspective can yield an adequate defense of democratic norms and institutions on the basis of such minimal epistemological resources.

Robert Sinclair

See also Epistemology; Instrumentalism; Truth

Further Readings

- Dewey, J. (1929). *The quest for certainty: A study of the relation of knowledge and action*. Minton, Balch and Co.
- Forster, P. D. (1998). Pragmatism, relativism and the critique of philosophy. *Metaphilosophy*, 29(1, 2), 58–78.
- Hookway, C. (1997). Logical principles and philosophical attitudes: Peirce's response to James' pragmatism. In R. A. Putnam (Ed.), *The Cambridge companion to William James*. Cambridge University Press.
- James, W. (1981). *Pragmatism*. Hackett Publishing. (Original work published 1907)
- Misak, C. (2013). *The American pragmatists*. Oxford University Press.
- Rorty, R. (1979). *Philosophy and the mirror of nature*. Princeton University Press.
- Talisie, R. B. (2014). Pragmatist political philosophy. *Philosophy Compass*, 9(2), 123–130. <https://doi.org/10.1111/phc3.12102>

PREDICTION

When faced with uncertainty, people from time immemorial have turned to prophets, seers, shamans, oracles, or clairvoyants to answer the urgent question “What will happen?” and the closely related question, “What shall we do?” Today, the responsibility to make predictions falls mainly to science. Scientists can inform about the future not because they possess some secret information about what happens next, but because science is able to reveal regularities and laws of nature. The term *prediction* refers to a variety of phenomena and tasks, including making weather forecasts, estimating the long-term effects of rising levels of greenhouse gas emissions, projecting election results or the effects of political reform, informing business planning, evaluating the market potentials of new products, or forecasting the outcome of scientific experiments. In contrast to *prophecy*, prediction is based on reliable method and equally reliable data: meteorological time series, annual reports, survey data, or the results of experimental investigation provide the basis for informed knowledge claims and allow for reasonably assuming regularities and continuities.

One of the most contested issues today is the influence of *big data* on the quality of scientific prediction. In recent years, improvement in information collection and information processing technologies has dramatically increased the availability of data. What remains to be seen is whether this development will lead to an exponential growth in scientific knowledge and contribute to the accuracy of scientific prediction, or succumb to the danger that relevant signals will get lost in the steadily growing noise of meaningless data.

In science, generating a hypothesis means to make a prediction. In general, predicting implies the making of well-founded statements about a specific phenomenon before it can be observed. This implies that a prediction must be more than an educated guess. Predicted outcomes must either be certain or (at least) be assumable with a sufficient degree of probability. Seemingly, predicting is not only about future knowledge, or knowing in advance, but also about matters not already known. In scientific contexts, this often refers to the capacity of a theory (or set of theories) to

predict specific outcomes, thus indicating the quality of a theory. Moreover, the significance and meaning of predictions in science differ between the opposing epistemic perspectives of deduction and induction. Furthermore, the relevance and reliability of scientific prognoses vary between probabilistic (stochastic/statistical) and full explanations (applying for all cases under all circumstances).

Finally, prediction has different meanings in the natural and in the social sciences. In the first case, predictions are correct epistemic claims on the basis of well-established and empirically confirmed theories. The inherent risk of making predictions in the *natural sciences* is fallacy. If the theoretical assumptions are incorrect and/or the data corrupted, the wrong conclusions will be drawn. In reverse, if a particular outcome predicted on the basis of a specific theory fails to materialize, then the quality of data (and observation) and/or the correctness of theoretical model must be reviewed.

In the *social sciences*, prediction is at the same time more difficult and much riskier. Social sciences deal with specific historical cases. Extrapolation from the data gathered and analyzed in a particular case is difficult because scientific laws in the social sciences differ from those found in the natural sciences. Moreover, because of the responsiveness of the very object of research (the social world) to scientific findings and claims, statements given in the social sciences can have an effect. Description can turn into prescription. In contrast to the natural sciences, the correctness of predictions in the social sciences not only depends on the proper application of theoretical models and the sound interpretation of empirical data but can result from the interplay between the scientists making the claim and their audience. This problem of potential *self-fulfilling prophecy* is a cardinal conceptual problem and a major obstacle to scientific prediction in the social sciences.

Explanation and Prediction From Deductive and Inductive Perspectives

Scientific explanation in general is based on regularities. This implies that a specific explanation

does not apply for a single event or phenomenon but can be hold true for all cases within the same category. Instead of explaining why a particular apple falls from a specific tree, the theory of gravity seeks to offer a credible explanation that applies to all apples (and all other objects with mass). The claim of modern science to uncover regularities and discover (universal) laws implies a claim to the capacity to predict certain events: If, on the basis of available data, a universal law can be derived, then this law should apply for all unobserved and future cases as well. *Because* until now all objects of the type *a* have under the circumstance *b* behaved in the way *c*, the same behavior can be predicted for this particular class of objects under the same circumstances.

In principle, two fundamentally different perspectives of scientific reasoning can be distinguished: deduction and induction. *Deduction* seeks to subsume specific cases under some general scientific laws and derive an explanation of a particular event by proving that it is an instance covered by a (set of) empirically proven regularities. *Induction*, in contrast, aims at coming to a general conclusion by observing single, individual instances. Both approaches are problematic from the perspective of the philosophy of science, especially with regard to their predictive power.

Deductive explanations have a strict logical form. It can be expressed employing the deductive-nomological model of explanation introduced by the German-American philosophers of science Carl Hempel and Paul Oppenheim. The model consists of the *explanans* (the general laws and particular facts that actually do the explanation) and the *explanandum* (the phenomenon that should be explained). The explanandum must be logically deducible from the explanans in order to satisfy the requirements of the deductive-nomological model of explanation. In the famous paper from 1948 titled "Studies in the Logic of Explanation," Hempel and Oppenheim argue in favor of a logical symmetry of scientific explanation and prediction. In general, the logical structure of the deductive-nomological model can be expressed as follows:

Explanans: The general laws of gravity tell that objects with mass fall to earth in a straight line. Apples are objects with a mass.

Explanandum: A particular apple falls to earth in a straight line.

According to Hempel the deductive-nomological model not only explains the trajectory of a falling apple but also allows for predicting apples to fall. In his view, explanation and prognosis are closely, logically interrelated: In the deductive-nomological model every explanation is a potential prognosis and vice versa. This is called the *thesis of structural identity*. Because in the deductive-nomological model the explanandum is logically implied by the explanans, the model can be used to formulate predictions with reference to the explanandum. The formalism of the deductive-nomological model of explanation and, in particular, the strict rules for scientific explanation render logically irrelevant the difference between (a) an explanation of cases already observed and (b) a prediction with regard to the outcomes of similar but unobserved cases. Because every scientific explanation requires the proof that a particular case (the explanandum) is covered by assignable laws and regularities, the explanation for similar cases can be given in advance of the actual observation. In reverse, the formulation of a prediction on the basis of established theories bears the risk of falsifying this particular theory or set of theories. The principle of *falsification* is a concept that demands scientific theories to make some definite predictions testable against empirical evidence. According to the Austrian philosopher of science Karl Popper, noted for his advocacy of this principle, risking falsification is both a necessary element of scientific progress and a reliable criterion for distinguishing science from pseudoscience. From this perspective, falsifying a scientific prediction proves the underlying theoretical concept wrong. Similar to Hempel and Oppenheim, he assumed a symmetry between scientific explanation and prediction. The value of prediction is primarily epistemic, for every explanation can in principle be expressed as a prediction. In this sense the deductive-nomological model of explanation offers a comprehensive but specialized framework for scientific prediction. Predictions can either be correct or not. If we can subsume a particular phenomenon under a class of cases for which deductive explanation have been found reasonable, then predictions can be made. Because the deductive-nomological model requires

the laws expressed in the explanans to cover all cases with specific conditions (e.g., there must be no exclusion from the universal law of gravity for any objects with a mass), the only difficulty is to determine whether a particular phenomenon is part of a larger class of cases. Probabilistic prediction ("it is highly likely that x ") is just as doubtful as prediction with regard to a phenomenon for which no general explanation is deducible.

In a sense, probabilistic explanation and prediction touch upon the core of inductive reasoning. In the case of predictions made on the basis of probability, correctness is a prime virtue, too. However, in drawing on the principles of inductive reasoning, they should be measured against the touchstone of accuracy. In contrast to explanation derived deductively, inductive reasoning starts with premises about cases already examined to draw conclusion about cases which have not yet been observed. Induction can be described as a bottom-up approach of scientific reasoning aimed at making generalizations on the basis of a finite number of observations. The observation of a number of falling apples n leads to the conclusion (prediction) that the next apple one might observe (n_{+1}) will behave in the same way. However, inductive reasoning poses some serious logical problems for making predictions. The cardinal problem of the inductive perspective is how to logically justify the assumption that the regularities or (presumed) causal relationships found in observed cases will persist into the future. How can particular outcomes of unobserved cases be predicted without referring to universal, covering laws?

From a logical point of view, prediction made on the basis of inductive reasoning is necessarily rendered preliminary and uncertain. This can be illustrated with reference to the *black swan problem*. Inductive reasoning rests on the assumption that one can arrive at general statements or a conclusion by making a series of singular existential statements based on observation. For instance, initial observation can reveal that a specific type of bird (a swan) is white. Subsequent observation of a number of swans may confirm this finding. As stated above, there is no logical argument for assuming that *all* swans are white since the discovery of a single black swan would falsify the conclusion. However, it may be possible that black swans are rare, so that the probability to correctly predict the color of swans is still very high. But how can

this belief in high probabilities of drawing the right conclusions be justified?

The 18th-century philosopher David Hume, noted for his work on inductive reasoning, was convinced that expectations about unknown cases are primarily determined by custom, understood as a subjective psychological disposition. Hume therefore offers a psychological explanation for the underlying mechanism behind empirical predictions: custom and habit. His explanation claims a lack of rational logical justification for this mechanism. The same is true for his *uniformity of nature* assumption, which develops this argumentation even further. It functions as a presupposition for inductive reasoning in general. The uniformity of nature assumption implies the fundamental belief that the natural world is guided by regularities (unobserved cases, which are in all relevant aspects similar to cases already examined, will be similar with regard to the conclusion). However, the truth of the uniformity of nature assumption itself cannot be proven. Ultimately, inductive reasoning for Hume rests on *good faith* rather than on logical principles. Therefore, inductive explanation and prediction always contains subjective elements (customs and degrees of conviction).

Although neither probabilistic prediction nor predictions based on inductive reasoning can attain definitive certainty, the plausibility (accuracy) of such predictions can be increased. Especially the Bayesian conception of subjective probability offers formal rules and procedures for enhancing the reliability of such predictions.

Prediction, Probability, and Bayes' Rule

Scientific prediction as well as scientific explanation can be based on probabilities (e.g., stochastic calculation and laws based on statistical connections). However, the concept of probability causes difficulties in science. First and foremost, the term itself can have different meanings. From a statistical point of view probability is closely linked to frequency. Scientific findings may indicate that in eight out of 10 cases y is the result of x . Therefore, the probability to predict y to be caused by x is 80%. The frequency interpretation of probability is unproblematic as long as it is based on actual (countable) evidence. It becomes more difficult

when based on proportions or relations alone. Given the vast number of solar systems in the universe, it is highly likely that a large number of planets similar to earth exist. Although it is probable that a number of those planets have suitable conditions for organic life, it is not possible to calculate the likelihood of extraterrestrial life or statistically predict its existence with the same degree of confidence as in the initial example.

Next to the frequency interpretation of probability are subjective assessments of probability. The subjective probability of a proposition expresses the degree of belief or credence of the researcher. This does not imply that subjective probability has to be arbitrary. *Subjective probability* rather refers to nonobserver-independent probability. In the scientific context, subjective probability is based on theoretical assumptions and existing evidence. Hypotheses, which include propositions based on subjective probability, refer to the presumed likelihood assigned to a hypothesis given the current state of knowledge. Accordingly subjective probabilities can—in line with the principles of inductive reasoning—be adjusted when new evidence becomes available. Using the Bayesian approach to probability can provide the need for rules how this adjustment could actually be made.

This approach, now among the most prominent interpretations of subjective probability, is named after 18th-century mathematician Thomas Bayes. He provided a formalized, machine-like procedure to evaluate the probability of a hypothesis. At the beginning the prior—that is, the subjective—probability of a particular hypothesis before any evidence is considered has to be defined. Additionally, the *posteriori* probability has to be set. *Posteriori probability* refers to the subjective probability immediately after relevant evidence came in. For calculating posteriori probability given defined prior probabilities, Bayes developed a formula for relating the (prior and posteriori) odds of event X_1 to X_2 conditioning on event Y (*Bayes' rule*). The relationship between the odds of X_1 and X_2 to Y is called the *Bayes factor* or *likelihood factor*. Sequential use of Bayes' rule (with posteriori probabilities of in the first round become priors in the second) offers a structured approach for systematically take evidence into account. Bayes interpretation of probability allows calculating changes of subjective probability in the light of (new) evidence.

The Bayesian approach provides standards for linking prior and posteriori probabilities and contributes to improving the quality of probability-based prediction. However, it is not possible to eliminate the subjective element altogether. It is therefore not possible to objectify subjective probabilities and degrees of belief in a way that the general certainty deficit of predictions based on induction could be solved.

Epistemic and Empirical Significance of Scientific Prediction

Prediction is generally understood as an informed statement about an unobserved phenomenon (x will occur). Against the background of fundamentally different approaches to scientific reasoning, and given difficulties of including probabilistic statements in scientific prediction, it is important to further distinguish between the epistemic and empirical significance of predictive propositions. As was demonstrated earlier, making correct predictions can be instrumental in testing the validity of general scientific laws. Therefore, in the perspective of deductive reasoning prediction is of epistemic significance, proving that the outcomes for a particular case (explanandum) were contained in the explanans. Making correct prediction also helps, when Bayes' rule is followed, to enhance the plausibility of subjective probability estimation, which forms the starting point of inductive reasoning. Making predictions from both perspectives has an important though somewhat different epistemic value. An important question, and ground for a continuing debate within the philosophy of science, is whether scientific theories must be able to make successful novel predictions. While this debate mainly focuses on questions of the nature of novelty (Should it be understood in a temporal sense only? Or should phenomena previously impossible or improbable to predict be included?), another line of the discussion is about confirming novel predictions, thus directly touching upon the empirical significance of prediction. From a deductive perspective, confirmation must take place within the logical structure of the deductive-nomological model. A true novel prediction eventually confirms that some known, general laws and regularities cover a particular, unobserved case.

Confirming prediction made on the basis of inductive reasoning and probabilistic statements means to enhance the probability to make correct statements by including relevant empirical evidence. Whereas the Bayesian rule provides procedures for the structural inclusion of such evidence, selecting and interpreting empirical findings pose some major methodological and epistemological challenges.

Prediction, Empirical Evidence, and Laplace's Demon

How to actually make empirical predictions? From the perspective of scientific determinism, the quality of a prediction depends on the complete knowledge of the underlying principles and laws as well as the inclusion of all relevant empirical data. When this knowledge is available, precise predictions can be made. The in-principle predictability of unobserved cases given complete knowledge of past and present states is frequently referred to as *Laplace's demon*, after 18th-century mathematician Pierre-Simon Laplace. As a thought experiment a supernatural entity (the demon) with complete knowledge (including the precise position and motion of every atom in the universe as well-secured information concerning the underlying laws) must be capable of correctly predicting all of their future conditions and so the state of the future. This argument is called *scientific determinism* and rests on the assumption that every state of affairs (including human action) is the inevitable consequence of previous events that are related in a regular manner. According to this view, the quality of predictions can be improved by adding further knowledge (empirical information or secured scientific knowledge about laws and regularities). Although scientific determinism has a strong appeal to the natural sciences, it is frequently criticized from the perspective of the philosophy of social science. One strand of criticism, which may be called the *pragmatic argument*, doubts the possibility of having complete knowledge of human action. For human action includes a large number or (potentially interacting) variables, such as psychological conditions, motivation, and belief systems. Therefore, it is impossible to gather sufficient knowledge. A second and perhaps more important strand of criticism aims at the deterministic belief in necessarily connected event

chains: Assume that one has complete knowledge including all empirical data and underlying laws relevant to a particular case at a given moment in time (t_n). Scientific determinism states that the conditions which led to the state t_n are the necessary consequences of previous events $t_{<n}$. From this perspective it would be possible to continue the chain of events to arrive at complete knowledge for a future state of affairs t_{n+1} . The question, and primary cause for criticism, is whether this is true for social developments. The state of affairs at t_{n+1} could be anticipated at t_n . Since the condition at t_{n+1} are (partially) the consequences of human action, and these actions are in turn premised on anticipations, the consequences of these anticipations (expressed in predictions about future states of the social world) manifest in the future state of affairs. This leaves unanswered the question whether the anticipations contain knowledge about inevitable developments or whether they are actually *causing* specific developments.

Prediction in Social Sciences: The Thomas Theorem and Self-Fulfilling Prophecy

The complexity and contingency of the social world are generally referred to in order to explain the relative predictive weakness of the social sciences. From the perspective of scientific determinism, it is particularly hard to arrive at sufficient (let alone complete) knowledge about social cases. The predictive power of the social sciences is seen as limited by insufficient methods of data ascertainment and analysis, and by qualitative differences between empirical evidence in the natural and the social sciences. In the natural sciences, empirical evidence provides information about what is the case. In the social sciences, evidence rather discloses what makes the case. This fundamental qualitative difference is summarized in the so-called *Thomas theorem*: If men define situations as real, they are real in their consequences. In the social world cases cannot be reduced to their objective features. Social action is always—and at times primarily—determined by the meaning of a particular situation. This poses a major problem for making predictions. For predictions in the social sciences not only run the risk of missing relevant knowledge but also affect

and alter the meaning of a given situation. The mechanisms and dangers behind this problem were illustrated in a classical text on self-fulfilling prophecies by sociologist Robert K. Merton, taking the example of a bank run: To issue a warning concerning the liquidity of a financial institution, economic research, including the analysis of accounting reports and balances, should reveal reliable data to justify the warning. Correct interpretation of available evidence allows anticipating financial turmoil and helps investors and savers to prevent losses. Based on scientific examination the warning can either be justified (supported by correctly interpreted evidence) or not. With reference to the Thomas theorem, Merton argues that scientific predictions (in this case the warning of bankruptcy) can be instrumental in causing the described effect. When investors and savers learn about the warning and rate it as trustworthy, they will act *as if* the warning were well founded. As rational agents they will pull out saving deposits and sell their stock at fire-sale prices, eventually causing the very liquidity problems they are trying to escape. With reference to the logical status of prediction in the social sciences, the relation between describing specific developments and creating or facilitating certain changes is highly problematic. This problem touches upon a major question and primary source of dissent in the social sciences. Can social sciences make predictions in the same way the natural sciences do?

Besides mentioning logical and epistemological restrictions, Karl Popper highlighted this question when he launched a general attack on historicism, formulating a fundamental criticism of the role of prediction and prophecy in the social sciences. *Historicism* refers to the doctrine that social sciences have the task of making historical prophecies (analogous to predictions in the natural sciences) and that these predictions are necessary to conduct politics in a rational way. Popper criticizes historicism and in particular the theoretical flagship of his time, Marxism, for being nonscientific and violating the methodological principles for making scientific predictions. Popper identifies two constitutive (sub) doctrines of historicism: (1) the historicist doctrine of social science, claiming that social sciences should make historic predictions about social developments (e.g., predicting

social revolutions) the same way that natural sciences make predictions about natural phenomena (e.g., solar eclipses); (2) the historicist doctrine of politics, which assigns the task of facilitating these developments to politics.

The first of these two doctrines is criticized for being methodologically untenable. Long-term *prophecies* (or predictions) can only be derived from conditional predictions (asserting that certain changes will be accompanied by others). Additionally, in line with the deductive-nomological perspective of scientific reasoning, in order to make certain (long-term) predictions, the superordinate system must be well isolated, stationary, and recurrent. Social systems, in contrast, are interdependent, dynamic, and at times produce original, unparalleled configurations. Predictions of social developments (e.g., revolutions) therefore are not only less certain but also cannot be derived from universal social laws. Because the objects of research (societies as a whole; specific groups like the middle class) are, according to Popper, more postulates of social theory than empirical objects; and since social change and societal development are not necessarily the result of deliberate design but can be the unintended consequences of intentional action, instead of making historical prophecies (which, in line with the second historicist doctrine, allow the provision of political roadmaps to ease inevitable developments), social sciences help to understand the possible consequences of human action. Hence, it is impossible to predict the necessary revolutionary development eventually leading to the overcoming of capitalist order with scientific certainty as Marxists intended to do. It is also impossible to formulate the necessary political accompanying measures to facilitate this development. Social sciences are nevertheless practically important. By analyzing the interplay of aspects of the social world (e.g., the consequences of normative concepts like *gender equality* for productivity and the formal organization of the labor market or economic principles of growth and the problem of externalization for environmental protection), social scientists are able to identify likely consequences, issue warnings of risks, and help political decision makers and societies in general to shape political programs. However, given the nature of societies and in particular the interplay of intentional action and unintended consequences, social sciences are limited to make probabilistic statements

(implementation of policy x is likely to have the effect y). Moreover, as the addressees of the findings of social sciences are able to interpret and interact with them, the danger that such statements will affect the probability of their being correct (turn out to be self-fulfilling prophecies) remains.

Empirical Prediction and Big Data

Big data (also referred to as *data science*) is a generic term for the increasingly available and evaluable large data sets. Improvements in data storage and processing capacities, along with new and advanced options in digital data collection, seem to be giving rise to Laplace's demon. Ranging from national economic statistics, communication reports, social media networks, and meteorological data to global positioning tracking, vast amounts of data on virtually all aspects of life have become available. With the empirical basis for prediction increasing, does this imply that the quality of scientific prediction increases accordingly? This question cannot easily be answered. The impact of information on the quality of scientific predictions depends, on one hand, on the quality of data itself and, on the other hand, on a proper understanding of the underlying principles and interdependencies. Whereas the first aspect refers to the crucial need to filter signals out of the rapidly swelling noise of context (i.e., to separate relevant from irrelevant information), the second aspect points toward the priority of useful criteria for this selection process.

Additional evidence can thus help to increase the plausibility of the subjective calculation of probabilities following Bayes' rule. Additionally, a growing number of observations increases the reliability of statistics based on frequencies. It is therefore plausible to expect advancement with regard to the correctness of empirical observation based on (relatively) simple theoretical models. To a certain degree, this is also true for empirical estimation based on highly complex data, as in weather forecasting. Although the fundamental relationship between key variables influencing the development of local weather phenomena are known, and additional data and time series improve the reliability of forecasting models, the system's complexity sets limits to the accuracy of the predictive efforts. Turning to complex

question where the underlying principles are not fully understood, the increase in available information actually can diminish the predictive power. With reliable criteria to distinguish *signal* from *noise* lacking, the risk of using irrelevant information rises considerably. Obviously, it is not the sheer amount of available information that determines the quality and reach of scientific prediction. Even with computational capacities increasing equally, the simple equation of more (information) = better (predictions) cannot be sustained. The increase in data availability conjures up a mentally impaired Laplace's demon unable to make sense of a vast amount of information. Without suitable theoretical assumptions for selecting relevant information, and lacking a working hypothesis about the connection and the phenomenon in question, the "garbage in, garbage out" principle will thwart possible improvements of increased information availability. This principle, borrowed from the field of computer sciences, refers to the fact that computational capacity cannot replace scientific theory development. For computers will process any (including nonsensical) data (garbage in) and will use programmed algorithms to produce respective (including nonsensical) output (garbage out). In consequence, big data will not automatically increase the predictive capacities of scientific inquiry. Improvement in data availability will only have an added value for scientific research and scientific prediction if the suitable theoretical models can be developed.

Dealing With Probability and Uncertainty: Forecast Models and Scenarios

What separates scientific predictions from prophecy? What are the principles of scientific forecasting? As stated above, distinct principles and perspectives of scientific reasoning assign specific epistemic and empirical values to predictions. Moreover, differences between the social and the natural realm reveal problems of scientific determinism and set limits to scientific predictions in both scope and accuracy. Finally, the increase in data availability and the advent of big data can enhance predictive power only when suitable scientific theories and methods are at hand.

As noted earlier in this entry, scientific predictions differ from prophecies by relying on models and methods. Reflecting the differences between natural and social sciences and distinct perspectives of scientific reasoning respectively, predictive methods can be categorized with reference to their primary objectives. From the perspective of deduction, predictions are made to test specific theoretical assumptions. The selection of methods for testing hypotheses and making predictions is determined by the respective epistemic goals. In the last four decades, an extensive set of quantitative and qualitative methods has emerged for making empirical predictions, complementing more traditional approaches of making predictions to test scientific hypotheses. Among the most prominent methods are (computational) modeling, especially relevant to climate science and meteorology, cross-impact estimation, which aims at calculating the co-occurrence and impact of events on the basis of conditional probabilities (predominantly used to predict the interplay of natural and societal change); scenario planning (a widely used technique, which arguably draws lessons from Popper's critique of historicism by avoiding *historic prophecies* and modeling plausible developments instead); and qualitative approaches like the Delphi methods (relying on expert opinion to improve subjective probability calculation).

A textbook example for replacing old (nonscientific) methods of predicting is weather forecasting. We need not go back to ancient or medieval times, when people attempted to predict the weather by interpreting signs and omens; rather, we can use the rise of quantitative weather prediction as a result of an increased demand for military meteorology since the 1920s. It was the demand to provide reliable weather forecasts for the emerging air forces during this time that triggered major changes in the science of meteorology. The production of weather maps, numerical modeling, and simulations were developed as a means to systematically include statistical evidence and improve the probability of occurrence (in accordance with Bayes' rule) in weather forecasts. The significance of computational models has further increased. In general, models are simplified representations of the world, which allow measuring and testing the influence of variables. They are particularly useful when the initial object of study (weather or climate systems) is too complex for

direct measurements. Because they allow testing the influence of specific variables repeatedly, they contribute to increasing the quality of estimating the subjective probability of a given variable to exert an influence. The aim of scientific modeling is to build sufficiently accurate models to allow useful predictions about the real world. Consequently, models can be good or bad.

Scenario development, in contrast, does not aim at providing an altogether accurate representation of the real world. Scenario development does not serve to eliminate uncertainty but rather to deal with it. Instead of feeding empirical information into a simplified model to measure the impact and interdependencies of single variables, scenario development aims at developing plausible narratives, laying out alternative ways of how the future might unfold. Accordingly, it is frequently used for strategic (economic and political) planning, as well as to explore the interplay of specific hazards and potential solutions (e.g., IPCC climate scenarios). Scenario development is frequently used to bridge the gap between findings and predictions (based for example on computational modeling) in the natural sciences and consequences for the social world. The scenario development process starts with the identification and assessment of *driving forces* relevant to a specific question. Additionally, time horizon and geographical scope (e.g., global carbon emissions, mitigation and adaption potentials by 2050) are specified. In keeping with Popper's critique of historicism, a number of scenarios (taking into account varying intensities or interrelationships of the driving forces) are being developed to present the possible connections and associated risks and to outline optional courses of action.

Computational modeling and scenario development are only some of many examples of the emerging field of futures research methodologies (a more comprehensive overview is outlined in the futures research methodologies handbook of the Millennium Project). Common to all methodological approaches is the need to conform to scientific standards. To make empirical predictions, methods need to be developed for the systematic selection and analysis of relevant data. To separate scientific prediction from prophecy, predictive methods

should consider the requirements and limitations of the opposing perspectives of scientific reasoning and the differing epistemic and empiric significance of predictions in science.

Prediction in Science, Prediction in Society

Prediction is an ambiguous term. This ambiguity derives from two sources: First, within science, predictions are of epistemic value, made to test hypotheses and be informative about the quality of theoretical models and assumptions. To make empiric predictions requires compliance with certain scientific rules (mainly for including and interpreting new evidence and calculating probabilities) and respecting the limits of predictive power (in particular in the social sciences). Second, these epistemological and methodological concerns at times conflict with the social demand for scientific advice. Science must of course make useful contributions to identify and deal with the challenges of the future. Predictions clearly are more than epistemic tools or empiric products of scientific research. They have effects in the social world, can provide desired or unwanted information, issue warnings, and inform about possible courses of action. As many observers have pointed out, it is crucial for science in general to maintain the distinction between scientific prediction based on theoretical assumptions and respecting clearly defined rules and methods, on the one hand, and prophecy the other. Prediction, as we have seen, is a scientific enterprise stemming from the desire to expand the frontiers of knowledge (either by testing hypotheses or by making accurate empirical predictions on the basis of theoretical concepts and models). Prophecy, in contrast, is a political enterprise making knowledge claims about unknown facts and the future without disclosing the processes and methods of how this particular knowledge was acquired.

Alexander Ruser

See also Accuracy; Epistemology; Modeling; Philosophy of Science

Further Readings

- Anheier, H. K., & Katz, H. (2009). Introducing futures research: Forecasting and scenarios. In H. K. Anheier, A. Kumar, J. A. Scholte, M. Kaldor, M. Glasius, & H. Seckinelgin (Eds.), *Global Civil Society Yearbook 2009: Poverty and activism*. SAGE Publications.
- Glenn, J. C., & Theodore, G. J. (2009). *Futures research methodology* (Version 3.0). The Millennium Project.
- Hempel, C. G., & Oppenheim, P. (1948). Studies in the logic of explanation. *Philosophy of Science*, 15(2), 135–175.
- Hyoiningen-Huene, P. (2013). *Systematicity. The nature of science*. Oxford University Press.
- Merton, R. K. (1948). The self-fulfilling prophecy. *The Antioch Review*, 8(2), 193–210.
- Orrell, D. (2007). *The future of everything: The science of prediction*. Thunder's Mouth Press.
- Popper, K. (2002). Prediction and prophecy in the social sciences. In *Conjectures and refutations. The growth of scientific knowledge*. Routledge. (Original work published 1963)
- Reichenbach, H. (1938). *Experience and prediction. An analysis of the foundations and the structure of knowledge*. University of Chicago Press.
- Rescher, N. (1998). *Predicting the future. An introduction to the theory of forecasting*. State University of New York Press.
- Sarewitz, D., Pielke, R. A., Jr., & Byerly, R., Jr. (2000). *Prediction. Science, decision making, and the future of nature*. Island Press.
- Schoen, E. L. (1991). The roles of prediction in science and religion. *Philosophy of Religion*, 29, 1–31.
- Silver, N. (2012). *The signal and the noise. The art and science of prediction*. Allen Lane, Penguin Books.

believe the world was created by God less than 10,000 years ago, defend what they call *creation science*, arguing that it should be taught as an alternative to mainstream science in public schools. Practitioners of homeopathy, the treatment of disease by minute doses of natural substances, claim that their field is a science, and (often) that it should be a form of publicly recognized (and funded) health care. Climate-change skeptics also produce evidence that purports to be scientific, arguing that such evidence undermines the consensus view of climate scientists. If we are to resist such claims, we would appear to need some way of distinguishing genuine science from what merely appears to be science. What better term for the latter than *pseudoscience*? This entry presents the definitions and conventions that support the distinction and explores the basis of the various critiques and defenses that have arisen within the communities of practice and scholarship in which these issues are contested.

Critics of the Distinction

Despite its apparent utility, the term *pseudoscience* has fallen upon hard times. Many historians and philosophers of science have questioned whether any principled and consistent distinction can be made between science and pseudoscience. The most radical of these critics suggests that the distinction resembles the border between two nations that share a common landmass. In itself it is arbitrary, reflecting nothing more than the conventions by which it is established and the interests of those who support them. On this view, one can describe the ways in which people *distinguish* between science and pseudoscience, but the distinction has no normative significance. It tells us something about the people labeling the practices in question, but nothing about the practices themselves. In particular, it does not correspond to a distinction between reliable and unreliable ways of attaining knowledge.

A second form of criticism takes a more moderate line. It does distinguish between reliable and unreliable ways of attaining knowledge but argues that these are not captured by the terms *science* and *pseudoscience*. This view was made famous in a 1983 essay by Larry Laudan who argued that

PSEUDOSCIENCE

The terms *pseudoscience* appears, at first sight, to be a useful one. Science is widely regarded as one of our most dependable ways of forming beliefs about the world. It may not answer all questions, but the questions it does answer, it answers with a high degree of precision and reliability. But there are those who claim the mantle of science—piggybacking, so to speak, on its reputation for reliability—whose credentials for doing so are dubious. Some young-earth creationists, who

attempts to distinguish science from pseudoscience have repeatedly failed. Early efforts at making such a distinction focused on method, assuming that there was a single scientific method. But this assumption has been shown to be false. Later attempts focused on the criterion of verifiability or falsifiability, the latter made popular by the work of the philosopher Karl Popper. A scientific theory was one that could be shown to be true, or (on the Popperian view) shown to be false. But many practices we may think of as pseudoscientific are eminently falsifiable. The efficacy of homeopathy, for instance, is testable by way of randomized trials. So this criterion also fails. Laudan is keen to distinguish between reliable and unreliable cognitive practices. But with regard to the term *pseudoscience*, his conclusion is very similar to that of the more radical critics: It is an empty term, of merely rhetorical significance. Since it does no useful work, it should be abandoned.

Defenses of the Distinction

So useful is the distinction, however, that it has not gone away. Recent attempts to defend the distinction between science and pseudoscience have involved two moves. A first move has been to broaden the scope of the discussion, from beliefs, propositions, or theories to the social practices that produce them. A second move has involved producing multifactorial accounts of what distinguishes science from pseudoscience. On this view, our knowledge-producing practices fall on a spectrum, from the clearly scientific at the one end to the clearly nonscientific at the other. When practices at the nonscientific end of the spectrum claim to be scientific, they may rightly be labeled as pseudoscientific. On this view, the distinction between science and pseudoscience is not identical with that between reliable and unreliable cognitive practices. Practices we term pseudoscientific may enjoy some degree of reliability. But they lack the particular kind of reliability associated with science.

Practice-Based Accounts

A first move in the defense of the category of pseudoscience is to clarify just what it is we are distinguishing. Early discussions tended to assume

that what had to be labeled as either scientific or pseudoscientific were propositions or theories or perhaps methods of inquiry. But more recent studies of science have made it clear that science is a social practice: a collective activity that has a goal and is governed by norms. The reliability of scientific inquiry is not simply a matter of methods of inquiry employed by individuals; it also depends on interaction with others. It relies on other people to provide a starting point for inquiry, to replicate experiments, and to review and interpret results. More precisely, science is a particular kind of practice: It is a variety of cognitive practice, one devoted to the attainment of knowledge. Cognitive practices do, of course, give rise to propositions and theories. But if we are to understand the difference between science and pseudoscience, we should not begin with propositions and theories. We should begin by examining the character of the *practices* from which propositions and theories emerge.

Not all writers who make this point do so by speaking of practices. Mario Bunge, for instance, speaks of *cognitive* or *epistemic fields*. But the elements that make up an epistemic field include a research community pursuing a certain line of inquiry. So this view also recognizes that science is a collective activity. By focusing on the structure of this collective activity, we can employ a wider range of criteria than would be available if we focused merely on propositions, theories, or individual methods of inquiry.

Multifactorial Accounts

This insight leads to the second move in the rehabilitation of the category of pseudoscience. Early efforts to distinguish science from pseudoscience sought to identify a single factor (such as falsifiability) that would mark off the two domains. But it has become clear that while each proposed criterion may have some value, no one criterion will suffice. For any one such criterion, there will be pseudosciences that meet it and fields that appear to be scientific that do not. (Young-earth creationism is falsifiable, for instance, whereas string theory arguably is not.) There *are* criteria that we can use in making such judgments, even if no single criterion among them is sufficient to allow us sharply to demarcate science from

pseudoscience. But they will allow us to say that a practice that purports to be scientific lacks a number of the characteristics of a science.

What kind of criteria should we employ in making such judgment? They should be criteria that have epistemic significance: features of scientific practice that account for its particular kind of reliability. We can zero in on such features by thinking, first of all, about the characteristics of an ideal epistemic community, one whose procedures are most likely to eliminate error and encourage the pursuit of truth. We can then identify further features that single out scientific practices from those of other epistemic communities.

Perhaps the most famous attempt to identify the features of an ideal epistemic community comes from the work of the sociologist Robert Merton. Merton identified four norms that he believed characteristic of scientific communities. The first is what he called *universalism*: Factors such as race, nationality, religion, gender, and class are considered as irrelevant in the assessment of practitioners' work. The second is what Merton calls *communism* (or, if you prefer, *communalism*): the idea that the results of inquiry are made freely available to all. Merton's third norm is *disinterestedness*. Although individual scientists may be far from disinterested, the scientific community has procedures to counter partiality. Finally, such a community is characterized by an *organized skepticism*: No ideas are exempt from critical scrutiny.

A more recent list emerges from the work of the philosopher Helen Longino who discusses the characteristics of what she calls *ideal epistemic communities*. Such communities, she argues, will not merely tolerate, but actively encourage, dissent. They will have forums for the assessment of claims and the criticism of evidence. They will have publicly recognized standards by which claims can be assessed. Finally, they will exhibit an equality of intellectual authority, in the sense that any qualified member of the community is entitled to contribute to the debate.

Criteria of this kind take us part of the way, but their scope is too broad. They do not help us to identify those practices we customarily describe as *scientific*. An ideal community of historical inquiry, for instance, would also have these features, yet we may not wish to regard it as scientific. (The English term *science* has a narrower scope than,

say, the German *Wissenschaft*, but it is this narrower use of the term in which we are interested here.) So some further criteria are needed. Since the success of the natural sciences in modern times owes much to their use of measurement, mathematical modeling, and experimentation, we might want to add these characteristics as well. Practices would count as scientific to the extent they not only meet the standards of an ideal epistemic community but also engage in mathematical modeling, measurement, and experiment.

Such conditions would appear to be jointly sufficient for calling something a science: a cognitive practice that had all of them would count as scientific. But they are not individually necessary conditions: A practice does not need to have them all to count as a science. Archaeology, for instance, could count as a science even if it does not employ experiment. (Many archaeologists today do engage in laboratory work, but there was a time when they did not.) A multifactorial account abandons the idea that we can find a set of *both* individually necessary *and* jointly sufficient conditions for identifying a science. It regards science as a *family resemblance* or *cluster* concept. The practices we regard as sciences can be loosely clustered by way of what Massimo Pugliucci calls *threads*: features that are common to two or more instances, although there may be no thread that links them all.

Practices on the Spectrum

The previous paragraph spoke of archaeology "counting as a science." But this phrase could be misleading to the extent that it suggests the existence of clearly defined categories. A multifactorial approach makes *scientificity* (to use an ugly term) a matter of degree rather than kind. On this view, our cognitive practices can be placed on a spectrum. At one end will be clearly nonscientific practices (such as journalism) and at the other end will be those that are undeniably scientific (such as particle physics). In between we will find practices such as historical writing, legal reasoning, *soft* sciences (such as sociology), and quasi- or proto-sciences (such as string theory).

The question then becomes: At what point on the spectrum do we draw a line and declare that a practice purporting to be a science is not? On a multifactorial view, there can be no single answer

to this question. (This is one of the reasons Laudan gives for rejecting the use of the category.) But this does not mean the term *pseudoscience* has no use at all.

There does exist one necessary condition for calling something a pseudoscience. It is that the practice in question must present itself as a science. Journalism is not a pseudoscience simply because it lacks many of the features of scientific practice, for journalists do not claim that their practice is a science. To call a practice a pseudoscience is a way of drawing attention to the lack of substance in a particular practice's claims to precision and reliability. The use of the term *pseudoscience* will always be a move in an argument, to be deployed in a particular context to indicate the hollowness of a practice's claims. But this does not mean that the application of the term is arbitrary: There will be features of the practice in question that warrant this appellation.

Nonscience, Bad Science, Science Fraud

The distinction between science and pseudoscience is related to, but not identical with, a number of other distinctions. A first is between science and nonscience. Science is a particularly reliable form of cognitive practice, insofar as its use of measurement and mathematics allows for precise predictions. But as we have already seen, there can be cognitive practices that are reliable sources of knowledge that do not use such methods and which would, therefore, not count as scientific. To declare, for instance that the writing of history is not a science is not to denigrate the practice of history; it is merely to note that it employs different methods. It employs, for instance, narrative explanations rather than theoretical ones, and backs up those explanations with documentary evidence rather than experiments. It is only when proponents of a practice lacking scientific credentials claim that their practice is scientific that the term *pseudoscience* may be applicable.

Pseudoscience is also to be distinguished from both bad science and science fraud. Both categories are most helpfully regarded as applicable to the activities of individuals rather than to the collective activities (the social practices) in which they are engaged. Bad science can involve factors

such as poor experimental design or faulty statistical analysis. Instances of bad science can be found within practices that are scientific, practices whose norms include sound experimental design and reliable statistical analysis. The problem with (merely) bad science is not a structural problem with the practice in which it occurs. It lies in the failure of individuals to follow the procedures that practice encourages. The idea of fraudulent science adds to this the idea of an intention to deceive. So like *bad science*, the label *fraudulent science* is most appropriately applied to the behavior of individual scientists. The term *pseudoscience*, by way of contrast, is not applicable (at least in the first instance) to the activities of individuals. It is a way of describing the falsity of the pretensions of a cognitive practice.

The Limits of the Concept

The idea of a pseudoscience remains an *essentially contested concept*, and the present entry has done no more than offer one way in which it has recently been defended. Critics will continue to insist that there is something unsatisfactory about so vague a demarcation criterion: one that ranks practices along a spectrum and chooses to describe those toward one end as *pseudoscientific* when it seems useful to do so. But it is sometimes useful to do so. The term *science* has not merely a descriptive role; it also functions as an honorific. It is used of practices to which we extend a presumption of reliability, whose results we are inclined to accept even when we cannot check them ourselves. So when a practice takes on the mantle of science, it is claiming for itself a certain authority. We need to know if a practice claiming to be a science has the features that would entitle it to this kind of respect.

What follows if it does not? On the view defended here, relatively little. If we admit that there are reliable cognitive practices that are not scientific, then it does not follow from the mere fact that a practice is pseudoscientific that it is unreliable. We can be rightly suspicious of practitioners who protest too loudly that theirs is a science. They may be trying to cover up weaknesses in the practice in question. But the existence of such weaknesses is not essential to the idea of a pseudoscience, as defined here. All that follows

from the application of the term *pseudoscience* is that a practice lacks the particular kind of reliability we have come to associate with modern science. That is a modest claim, but one sometimes worth making.

Gregory W. Dawes

See also Authority; Climate Science; Complementary Medicine; Falsifiability; Philosophy of Science; Social Construction of Scientific Knowledge

Further Readings

- Bunge, M. (1991). What is science? Does it matter to distinguish it from pseudoscience? A reply to my commentators. *New Ideas in Psychology*, 9(2), 245–283.
- Dawes, G. W. (2018). Identifying pseudoscience: A social process criterion. *Zeitschrift für allgemeine Wissenschaftstheorie/Journal for General Philosophy of Science*, 49(3), 283–298.
- Hansson, S. O. (2017). Science and pseudo-science. In E. N. Zalta (Ed.), *The Stanford encyclopedia of philosophy* (Summer 2017 Edition) <https://plato.stanford.edu/archives/sum2017/entries/pseudoscience/>
- Hoyningen-Huene, P. (2013). *Systematicity: The nature of science*. Oxford University Press.
- Hull, D. L. (1988). *Science as a process: An evolutionary account of the social and conceptual development of science*. University of Chicago Press.
- Koertge, N. (2013). Belief buddies versus critical communities. In M. Pigliucci & M. Boudry (Eds.), *Philosophy of pseudoscience: Reconsidering the demarcation problem* (pp. 165–180). University of Chicago Press.
- Laudan, L. (1983). The demise of the demarcation problem. In R. S. Cohen & L. Laudan (Eds.), *Physics, philosophy and psychoanalysis: Essays in honor of Adolf Grünbaum* (Boston Studies in the Philosophy of Science 76, pp. 111–127). D. Reidel.
- Longino, H. E. (1994). The fate of knowledge in social theories of science. In F. F. Schmitt (Ed.), *Socializing epistemology: The social dimensions of knowledge* (Studies in Epistemology and Cognitive Theory, pp. 135–157). Rowman & Littlefield.
- Mahner, M. (2013). Science and pseudoscience: How to demarcate after the (alleged) demise of the demarcation problem. In M. Pigliucci & M. Boudry (Eds.), *Philosophy of pseudoscience: Reconsidering the demarcation problem* (pp. 29–43). University of Chicago Press.
- Merton, R. K. (1973). The normative structure of science (1942). In N. W. Storer (Ed.), *The sociology of science: Theoretical and empirical investigations* (pp. 267–278). University of Chicago Press.
- Pigliucci, M. (2013). The demarcation problem: A (belated) response to Laudan. In M. Pigliucci & M. Boudry (Eds.), *Philosophy of pseudoscience: Reconsidering the demarcation problem* (pp. 9–28). University of Chicago Press.
- Woit, P. (2006). *Not even wrong: The failure of string theory and the continuing challenge to unify the laws of physics*. Jonathan Cape.

PUNCTUATED EQUILIBRIUM

Punctuated equilibrium is the theory that one of the dominant modes of evolutionary change involves large, rapid changes in the morphology of a lineage, followed by evolutionarily long-term morphological stasis. It was introduced in 1972 to demonstrate the importance of theorizing in paleontology, as well as to broaden our understanding of the evolutionary process. This entry describes the historical foundations of this effort, its effects on the discipline of paleontology, and its broader significance for the study of evolution.

The Theoretical Significance of the Fossil Record

By the 1940s, paleontologists had arrived at a crossroad. On the one hand, their discipline had become increasingly involved in the petroleum industry, aiding in the search for oil. On the other hand, paleontology had been sidelined in the development of evolutionary theory as the powerful theory of Mendelian genetics merged with the even more powerful theory of Darwinian evolution, resulting in what is known today as the neo-Darwinian *Modern Synthesis*. As paleontologists stared out from this crossroad, it looked as if their future as an autonomous scientific discipline was in jeopardy. The developing theory of neo-Darwinian evolution seemed not to have any particular need for whatever information fossils could provide, and greater involvement with the petroleum industry would only steer them further away from biological science.

Paleontology's skies partially cleared in 1944, with the publication of George Gaylord Simpson's *Tempo and Mode in Evolution*. Of the many distinctive contributions of the book, Simpson's most important, lasting achievement was to show that patterns of long-term evolutionary change in the fossil record accorded with biologists' newly achieved understanding of how the composition of a population changes adaptively in response to environmental pressure. Simpson argued that the importance of this match between macro- and microevolutionary patterns lay not in the fact that it provided further evidence that Darwinian theory was correct. Rather, what was really significant about it was that it showed that the fossil record was a repository of genuine evolutionary patterns, perhaps only *some* of which reflected the influence of Darwinian mechanisms. A fully mature evolutionary theory, he argued, would have to be able to account for all patterns in the fossil record patterns, *whether or not* they confirmed expectations that were specifically Darwinian.

Simpson's insight was a lifeline to paleontology as it struggled to find its way forward, for it outlined a distinctive role for the study of fossil patterns in the future development of evolutionary theory. The theory of evolution could not be considered complete so long as the fossil record remained unaccounted for. Since paleontologists alone were competent to study it, there was an opportunity for them to contribute something to evolutionary theory which fell exclusively under the purview of their discipline. Indeed, Simpson himself had described an apparently non-Darwinian evolutionary mechanism in *Tempo and Mode*, which he called *quantum evolution*, in which large leaps across morphospace seem to occur over a timescale that made little sense from a traditional Darwinian perspective. Despite the promise that it held for paleontology's disciplinary ambitions, Simpson would eschew its non-Darwinian status in subsequent work, attributing it instead to allopatric speciation, which was by then part of the neo-Darwinian orthodoxy.

That transformative potential was not lost on Stephen Jay Gould, who, from his first days as a Columbia graduate student in 1965, saw that the survival of paleontology as an autonomous science rested solely on its ability to make contributions to evolutionary theory that could only be

derived from fossil patterns. Gould would develop Simpson's perspective on the evolutionary significance of the fossil record into lifelong fixation on the need for paleontology to articulate laws of nature that governed large-scale evolutionary change and which therefore formed an essential component of a complete theory of evolution. Gould's preferred vehicle for that fixation would be the theory of *punctuated equilibrium*.

Origins of the Theory of Punctuated Equilibrium

Interestingly, the theory of punctuated equilibrium began not with Gould but with another Columbia graduate student, Niles Eldredge. In 1971, Eldredge published a paper that would help make theorizing a permanent fixture in paleontology. This paper made two significant observations. The first was that the received view on how speciation occurs—the allopatric model—was not obviously consistent with the dominant mode of evolutionary change, in which lineages gradually transform through adaptation to their environments. Allopatric speciation, by contrast, involves relatively rapid morphological divergence in which genetic drift is assumed to play a central role. Second, Eldredge described the curious tendency of a certain trilobite lineage to undergo geologically rapid change (consistent with allopatry) but then persist essentially unchanged for millions of years, a feature that he attributed to stabilizing selection.

Gould seems to have seen in this work a singular opportunity to challenge the neo-Darwinian orthodoxy. The year before Eldredge's paper was published, Gould had been invited to present a paper on theoretical models of speciation and their role in paleontology. Not knowing much about the topic himself, Gould asked Eldredge to join him as a coauthor. The resulting paper, "Punctuated equilibria: An alternative to phyletic gradualism," reiterated Eldredge's emphases on both the illness of fit between the allopatric model and the received view on evolutionary change, and on the surprising long-term stability of several fossil lineages. Whereas Eldredge's original paper had attributed this stasis to stabilizing selection, however, Eldredge and Gould's (1972) "Punctuated equilibria" would make the

unorthodox claim that stasis was a consequence of allopatric speciation.

This seemingly minor change in view held important implications for paleontology's ability to make a distinctive contribution to evolutionary theory. Paleontology could not claim that allopatric speciation has a phenomenon uniquely disclosed by the fossil record. Nor would its theoretical ambitions be facilitated by the embrace of stabilizing selection as an explanation for stasis, since such an explanation would have been consistent with the generally Darwinian understanding of evolution. However, by linking stasis in the fossil record to the non-Darwinian nature of allopatric speciation and treating them as two stages of a single causal process—the *punctuated* morphological change symptomatic of allopatry, which causally gave way to an *equilibrium* stage lasting millions of years—paleontology could assert exclusive access to a fossil pattern not predicted by Darwinian theory.

The Transformative Role of the Theory of Punctuated Equilibrium

In this way, the theory of punctuated equilibrium was developed more out of paleontology's need for an original theoretical framework that could distinguish it from Darwinian microevolutionary theory than from a pressing need to solve an evolutionary puzzle. Neither allopatric speciation nor morphological stasis was considered to be a puzzle in 1972. Eldredge and Gould reinterpreted the fossil record (as well as the allopatric model of speciation) in order to manufacture a theoretical puzzle that only paleontologists could solve. Viewed from this perspective, the theory of punctuated equilibrium did not succeed in providing the "alternative to phyletic gradualism" that would have helped to clear some space for paleontology in the ecology of evolutionary disciplines. Orthodox Darwinian critics of the theory claimed that both *punctuation* and *equilibrium* could be understood using traditional and more conservative population genetic theory, criticisms that Eldredge and Gould would eventually accept. The received view is that what appears to be stasis in the fossil record is in reality the failure of the many short-lived instances of evolutionary change to leave fossil traces, and that the reason morphological changes appear to occur

primarily during speciation events is because speciation is the only kind of evolutionary change that really sticks.

Looking at the theory of punctuated equilibrium in a broader historical and disciplinary context, however, it made an undeniably lasting impression on evolutionary thinking and on the fortunes of paleontology in particular. For Eldredge and Gould had compelled the community of evolutionary biologists to finally heed the warning that had been issued by Simpson a generation earlier: A complete understanding of the evolutionary process cannot be derived from studying changes in gene frequencies over short periods of time. There are evolutionary phenomena that can only be observed on the longer timescales represented in the fossil record. Stasis is one of them. That it turns out to have a perfectly serviceable Darwinian explanation is immaterial to the fact that, had it not been for paleontology, we would not have known that the evolutionary process is compatible with millions of years of morphological stability.

The acknowledgment of *stasis as data* and of paleontology's necessary role in alerting us to it, combined with Gould's consistent and high-profile advocacy for the theory of punctuated equilibrium, form a major part of the explanation for why paleontology would grow into an increasingly theoretical discipline after the early 1970s. After that point, it could no longer be maintained that paleontology was merely a form of biological stamp collecting. Paleontologists were disclosing new evolutionary phenomena and were developing novel explanatory theories to account for them, some of which were significant departures from traditional Darwinian thinking.

Gould was involved in several independent efforts to advance the disciplinary status of paleontology; each of them focused on making a specifically paleontological contribution to our understanding of the nature of evolutionary change. Although none of his specific proposals ever enjoyed widespread uptake, he understood perhaps better than any other member of his generation that the future of the discipline hung in the balance, and that its very survival as a science would be determined by the extent to which evolutionary theorizing could be elevated to a central component of paleontological practice. His efforts

to introduce a culture of theorizing into paleontology (in which he was joined early on by others such as David Raup and Thomas Schopf) were successful not only in transforming the discipline but also in helping to recapture the essential breadth of evolutionary theorizing that Darwin himself had exhibited in the *Origin*.

Chris Haufe

See also Biology, Evolutionary; Cell Theory; Genetic Drift

Further Readings

- Eldredge, N. (1971). The allopatric model and phylogeny in Paleozoic invertebrates. *Evolution*, 25, 156–167.
- Eldredge, N., & Gould, S. J. (1972). Punctuated equilibria: An alternative to phyletic gradualism. In T. J. M. Schopf (Ed.), *Models in paleobiology* (pp. 82–115). Cooper & Co.
- Eldredge, N., Thompson, J. N., Brakefield, P. M., Gavrillets, S., Jablonski, D., Jackson, J. B. C., Lenski, R. E., Lieberman, B. S., McPeck, M. A., & Miller, W., III. (2005). The dynamics of evolutionary stasis. *Paleobiology*, 31(2), 133–145.
- Futuyma, D. (1987). On the role of species in anagenesis. *The American Naturalist*, 130(3), 465–473.
- Gould, S. J. (2002). *The structure of evolutionary theory*. Harvard University Press.
- Haufe, C. (2015). Gould's laws. *Philosophy of Science*, 82(1), 1–20.
- Sepkoski, D. (2012). *Rereading the fossil record*. University of Chicago Press.

R

RATIONAL MECHANICS

This locution was introduced by Sir Isaac Newton in the Preface to his *Philosophiae Naturalis Principia Mathematica* (*Mathematical Principles of Natural Philosophy*):

The ancients divided mechanics into two parts: the rational, which proceeds rigorously through demonstrations, and the practical. [. . .] In this sense, rational mechanics will be the science, expressed in exact propositions and demonstrations, of the motions that result from any forces whatever and of the forces that are required for any motions whatever. (1867, p. xiv; first English translation 1729)

Newton argued that rational mechanics is a branch of mathematics and went as far as to state that geometry is part of mechanics. Modern scientists are likely to disagree with Newton on this second point; to them, mathematized mechanics is geometry plus motion, time, mass, and force.

The new terminology was widely accepted on the Continent but was less common in Great Britain, perhaps because of the British distrust, in the period 1750 to 1820, of highly theoretical studies. By the 1950s, in English-speaking countries, the elegant and precise expression *rational mechanics* was considered archaic. It was supplanted by *theoretical mechanics*, which excludes rational theories of materials, and *classical mechanics*, which is too generic. The following sections trace

the history of rational mechanics from its origins in classical Greece to the contributions of Renaissance thinkers such as Galileo, Huygens, and Descartes and continuing with Newton and Enlightenment figures such as Euler, Bernoulli, D'Alembert, and Lagrange, and the work of their successors in the 19th and 20th centuries. The entry continues with an account of the theoretical framework built on the foundation of Newton's achievement, followed by a description of subsequent efforts to axiomatize the concepts underlying the fundamental principles of rational mechanics. The entry concludes with an explanation of the mechanist philosophies and models of the universe that held sway until the early 20th century, when the advent of thermodynamics, electromagnetism, and the theories of relativity and quantum physics would challenge the basis of the mechanical view of the world.

History

As Newton said, the substance of the theory goes back to the ancient Greece. In the 3rd century B.C.E., Archimedes had already rigorously formalized the law of the lever and the principles of hydrostatics in works that survive today. For two millennia, Archimedean mechanics set exacting standards of excellence.

By the early 17th century, scientists were ready to take up the challenge. In 1638, Galileo Galilei published a short Latin treatise entitled *De Motu* (*On Motion*), in which he established in purely Euclidean fashion the laws of uniform and

uniformly accelerated motion. His pupil Evangelista Torricelli gave a better formulation of this question in his *On the Motion of Heavy Bodies* (1644). That very same year, René Descartes gave a precise set of (wrong) rules for impact. This early period of discovery reached a pinnacle with the publication of Christiaan Huygens's *The Pendulum Clock: Or Geometrical Demonstrations Concerning the Motion of Pendula* (1673), today mainly cited for the theorems on pendular motion and centrifugal forces, which became the model for all subsequent rational enquiries into natural philosophy.

The publication in 1687 of Newton's *Principles* raised mechanics from infancy to adulthood. Modern readers expecting a textbook on dynamics will be dismayed; this is a product of Baroque times, to be savored like the music of Bach. After two introductory sections on *Definitions* (i.e., space, time, and motion) and the *Axioms of motion* (the three Newtonian laws of dynamics), the treatise is divided into three books: The first discusses the motion of a point attracted to one or two fixed centers by central forces; the second derives approximate formulas for bodies moving in a resisting fluid; the third contains the law of gravitation, from the first principles to the irregularities of moon motion.

The *Principles* is a concise and intensely mathematical work, and rare is the reader, then or now, who can master its content in a lifetime. It took the world a generation or two to catch up with Newton's ideas. Then, a handful of brilliant mathematicians, all born on the Continent at the beginning of the 18th century, who as youths had absorbed Newton's and Gottfried Wilhelm Leibniz's writings, were able to create new theories, expressed in terms of Leibnizian calculus, out of the Newtonian foundations.

It all happened in less than 30 years. In 1736, Leonhard Euler published the two-volume treatise *Mechanics, Or the Science of Motion, Analytically Expounded*, where the dynamics of a single particle appears as a branch of three-dimensional analytic geometry. His friend Daniel Bernoulli, in the *Hydrodynamics* (1738), transformed hydraulic engineering into a scientific theory based on what we would now call the conservation of mechanical energy (Bernoulli's theorem). In 1743, the mathematician and philosopher Jean D'Alembert, today

mainly famous as an exponent of the Enlightenment, reduced dynamics to statics by considering the equilibrium of those components of the applied forces which are destroyed by the constraints (D'Alembert's principle). Alexis Clairaut, in his *Theory of The Figure of the Earth* (1743), reduced fluid statics to a few equations.

The most remarkable exponent of this group is undoubtedly Euler. In a series of brilliant papers published from the 1740s onward, he single-handedly created fluid dynamics, linear elasticity, and the theory of rigid bodies in a form not very different from that encountered in our textbooks. By forging the field viewpoint in mechanics, Euler shaped the development of mathematical physics for the next two centuries.

An approach to rational mechanics different from Euler's was initiated by the Italian-French mathematician Joseph-Louis Lagrange. His *Analytical Mechanics* (1788) was an attempt to base all of mechanics on the general laws of statics, supplemented by the principle that reversed accelerations are forces per unit mass. Lagrangian mechanics is completely analytic, without any trace of geometry or physical intuition. It is one of the best examples of an uncompromising attempt at a rigorous foundational program in science.

By about 1820, most of the elementary theory had already been established and expounded in textbooks. At the same time, higher research advanced at a rapid pace. The engineer-mathematician H. Navier gave the general equations of linear elasticity (1821). Augustin-Louis Cauchy discovered the stress tensor, which expresses the properties of the surface forces in a material continuum (1829). William Rowan Hamilton, starting from Lagrangian mechanics, created a new form of the equations of motion of a system of particles, one which unveils a fascinating dualism between dynamics and geometrical optics (1834). Gaspard-Gustave Coriolis discovered the equations of relative motion (1835).

In the 20th century, with the arrival of elementary particles and the theories of relativity and quantum physics, classical mechanics was exiled from the mainstream of science. Yet, unbeknownst to physicists, the theory continued to grow. Prominent among the new developments was, after the Second World War, an impressive expansion of continuum mechanics, mostly under the energetic

leadership of Ronald Rivlin in Great Britain and Clifford Ambrose Truesdell III in the United States. Besides discovering a host of new results, Truesdell and his students formalized *rational continuum mechanics*, setting it firmly in the old tradition.

Theory

The general framework of rational mechanics was artistically polished in all its myriad detail during the 19th century. As a result of this activity, not much has changed in the exposition of its foundations since the appearance of a number of influential treatises in the early decades of the 20th century. What follows is a schematic summary of the key concepts.

First comes *statics*, the theory of equilibrium, and *kinematics*, the description of motion, then *dynamics*, the study of forces and motion. While kinematics is essentially the Euclidean geometry of moving bodies, statics involves two new physical entities: mass and force. *Mass* is the quantity of matter in an object and is measured by its inertia, that is, its resistance to a change of velocity. *Force* is a measure of the interactions between objects or between objects and fields. The product of mass and velocity is the *momentum* of a moving object, which is a measure of its quantity of motion. Force, velocity, and momentum are vectors.

The fundamental laws of dynamics are Newton's three laws of motion. Despite the strength of Newton's polished style, they are best understood in a slightly modernized form.

Lex I: In the absence of forces, a particle moves with constant velocity.

Lex II: The rate of change of the momentum of a particle is equal to the impressed force.

Lex III: For every force from Particle A on Particle B, there is an equal and opposite force from B to A.

These axioms apply to mass points; in other words, particles or geometrical points endowed with mass and motion.

The first law defines a class of reference frames, the inertial frames. Any two inertial frames move relative to one another with uniform velocity. Newton's laws only apply to motions relative to an inertial reference frame. The first law also defines Newtonian time.

When describing systems of particles, three additional quantities become important. The total kinetic energy is half the sum of the products of masses and their squared velocities. The work done by a force is the product of the magnitude of the force and the distance moved in the direction of the force. The *moment* of a vector (e.g., force or momentum) is the cross-product of the position vector with that vector. In systems, we must also distinguish between *internal* forces, acting between any two particles in the system, and *external* forces, exerted by agents outside of the system.

These definitions being provided, we obtain from Newton's laws three general principles which apply to all systems of particles. The *principle of momentum* establishes that the rate of change of the total momentum is equal to the sum of the external forces. According to the *principle of moment of momentum*, the rate of change of the total moment of momentum about any fixed origin is equal to the sum of the moments of the external forces about that origin. Finally, the *work-energy principle* states that, in any displacement of a body, the change in the kinetic energy is equal to the work done by all forces, internal and external.

From these principles, the *conservation laws* are immediately obtained. If the sum of the external forces is zero, then the momentum is conserved; similarly, if the sum of the moments of the external forces is zero, then the moment of momentum is conserved. When the forces are derived from a potential function, then mechanical energy is conserved too.

Next comes the theory of *rigid bodies*. The kinematics of rigid bodies is the purely geometrical theory of groups of isometric motions in space. Combining these kinematic results with the principle of moment of momentum yields Euler's equations, which express the properties of rotary inertia and explain the counterintuitive behavior of massive, swiftly rotating objects.

Besides this central core, rational mechanics has many subdisciplines. Among the most prominent ones, in no particular order, are analytical mechanics, statistical mechanics, celestial mechanics, general continuum mechanics, elasticity, fluid mechanics, and acoustics. Each of them has its own set of rules and is partially independent of the others. Unfortunately, even a cursory survey of

their characteristics would far exceed the scope of this article.

Axiomatics

After the appearance of Newton's *Principles*, the further development of rational mechanics raised a number of foundational problems. The evidence for or against the existence of absolute space and time, the physical meaning of inertia, the equality of inertial and gravitational mass, the logical status of the concept of force, and the degree of generality of the different principles of mechanics worried scientists for nearly two centuries. Although the relevance of these problems has greatly diminished since the advent of modern physics, a tinge of insecurity still encircles the edge of pedagogical presentations.

One way to overcome this situation would be to produce an axiomatization of those principles which intuitively hold for classical space, time, matter, and motion. This would amount to solving the heart of the Sixth Hilbert Problem, proposed by David Hilbert in 1900. Hilbert asked for an axiomatization of mathematical physics, starting from mechanics.

There have been three main attempts at a solution of this part of the Sixth Hilbert Problem, greatly differing in aims and methods. In 1908, the mathematician Georg Hamel, a student of Hilbert, taking inspiration from Hilbert's axiomatization of Euclidean geometry, introduced groups of postulates for the mechanics of extended bodies. About 50 years later, in a different scientific climate, the logicians John C. C. McKinsey, A. C. Sugar, and Patrick Suppes strove for a Bourbakist approach to point mechanics. Somewhat later the engineer-turned-mathematician Walter Noll sought instead for a conceptual framework general enough to include the then new theories of continua. Variations on these themes have been put forward over the years. Regrettably, as late as the second decade of the 21st century, there is no general agreement as to the respective merits of these competing views.

Mechanicism

It is only natural to think of the world as a complex machine. This doctrine, called *mechanicism*,

was an essential feature of the Scientific Revolution. It was originally sparked by the rediscovery of the ancient atomism during the Renaissance and continued to develop over the centuries. Although far from providing a correct interpretation of the universe, its different incarnations gave direction to the research of physicists.

A full-blown mechanistic model of the world was developed by René Descartes in 1644. According to Descartes, the cosmos is filled with matter. The essence of matter is extension. Vortices and currents of matter are the causes of gravity. Animals are pure machines. An external God is the ultimate mover of these fluxes of matter.

Descartes's dynamical universe broke completely with the Aristotelian tradition and marked most of the development of physics up to the 1720s. It gave us the principle of inertia, the conservation of momentum, a theory of light, and an explanation of planetary motion but was not grounded on experiments and could not predict new phenomena. In the end, it was denounced as fiction by intellectuals as diverse as Huygens and Voltaire. After an academic struggle that lasted for decades, around 1740, Descartes's philosophy was superseded by Newtonian science.

Newton's mechanicism owed nothing to Descartes's vortices. It was essentially a theory of forces, grasped by experiments and described through mathematics. This abstract structure is an attribute of an otherwise incomprehensible God. Newton also believed in the existence of atoms and of "a very subtle spirit pervading gross bodies and lying hidden in them," but since he could not prove his ideas with facts, he did not elaborate further.

Several mechanistic theories of nature arose during the 19th century. Between 1800 and 1830, Pierre-Simon Laplace, Siméon Denis Poisson, and Claude-Louis Navier attempted to establish a new science, called *mécanique physique*, which aimed at explaining capillarity, elasticity, and fluid dynamics by considering hypothetical intermolecular forces. Another unified theory of the world was *energetism*, especially advocated by the chemical physicist Wilhelm Ostwald in the 1890s, which purported to derive all of physics from the principle of conservation of energy. Also of importance were the various mechanical models of the *luminiferous ether* propounded by British

physicists from the 1860s onward. Clearly, every one of these philosophies contained a grain of truth but was far from being entirely correct.

Mechanicist philosophies reached an end in the late 19th century, with the advent of the new sciences of thermodynamics and electromagnetism. In his influential book *The Science of Mechanics* (1883), Ernst Mach remarked that the older knowledge is not necessarily the foundation of more recent theories. A few years later, Albert Einstein's special theory of relativity (1905) gave the death blow to all attempts at purely mechanical explanations of the world.

Sandro Caparrini

See also Laws (Scientific); Mathematics, Enlightenment; Physical Theory; Physics, 19th Century; Physics, Antiquity; Physics, Enlightenment

Further Readings

- Banach, S. (1951). *Mechanics*. (E. J. Scott, Trans.). Monografie Matematyczne Tom xxiv (Warsaw). Stechert-Hafner.
- Caparrini, S., & Fraser, C. (2013). Mechanics in the eighteenth century. In J. Buchwald & R. Fox (Eds.), *The Oxford handbook of the history of physics* (pp. 385–405). Oxford University Press.
- Dijksterhuis, E. J. (1986). *The mechanization of the world picture: Pythagoras to Newton* (C. Dikshoorn, Trans.). Princeton University Press. (Original work published 1950)
- Dugas, R. (1955). *A history of mechanics* (J. R. Maddox, Trans.). Editions du Griffon.
- Gurtin, M. E. (1981). *An introduction to continuum mechanics*. Academic Press.
- Mach, E. (2013). *The science of mechanics: A critical and historical exposition of its principles* (T. J. McCormack, Trans.). Cambridge University Press. (Original work published 1883)
- McKinsey, J. C. C., Sugar, A. C., & Suppes, P. (1953). Axiomatic foundations of classical particle mechanics. *Indiana University Mathematics Journal*, 2(2), 253–272.
- Papastavridis, J. G. (2014). *Analytical mechanics: A comprehensive treatise on the dynamics of constrained systems*. World Scientific.
- Synge, J. L., & Griffith, B. A. (1959). *Principles of mechanics*. McGraw-Hill.
- Truesdell, C. (1968). *Essays in the history of mechanics*. Springer-Verlag.

RATIONALITY

There are numerous models of *rationality*, the term often used interchangeably with *reason* to describe logical mental operations of human beings. The notion of rationality is at the core of epistemology and formal logic. Throughout this essay, *rationality* is an umbrella term to describe various types and functions of focused human intellectual activity as well as the instrumental use of thought and thought processes to carry out epistemological endeavors and achieve pragmatic goals. As generally conceived of in the Western tradition, especially since the 17th century, human rationality works by means of cerebral operations in, upon, and through the natural, material world of existence as a function of the human capacity to reason. Consciousness and self-consciousness, perception and logic, ideas and information, thoughts and thinking, and similar immaterial actualities are seen as by-products of particular areas of the physical brain in communion with faculties of the mind. In this essentially Western worldview, *knowledge* has been understood to be rooted primarily in processes of rationality, rather than in emotion, intuition, or the *received knowledge* of tradition or religion. Contemporary philosophy and epistemology, however, are paying much attention to the latter categories—how to understand them as kinds of knowledge and bringing them conceptually under the umbrella of what is described here as rationality.

Recent developments in epistemology, psychology, neuroscience, and computer science have spurred further interest in resolving the *mind-body* conundrum that is the offspring of René Descartes' *cogito ergo sum* (I think, therefore I am). Neuroscientists, psychologists, social scientists, and philosophers increasingly focus on the systematicity of thought and cognitive domains. *Mind-brain identity theory* and similar models emerge from these disciplines, all of which attempt to decipher connections among mental states and processes and interactions of the physical brain with the material world. Theories of rationality are often divided into two broad categories: *normative* or *theoretical* (what rationality is or might be) and *descriptive* or *practical* (what rationality does or should do). Similarly to the exposition of

rationality theories, theories of consciousness and representation continue to evolve in great depth and detail; owing to limitations of space, however, that discussion is not included in this overview.

Theories of rationality—and by extension, dicta as to who and what is *rational*—are instrumental in fields of politics and government, law, economics, political science, sociology, and cognitive psychology. The intensity of the centuries-long debate over rationality, its characteristics, and its proper application is precipitated at least in part by the close relationship between claims to rationality and claims to authority and power. The following discussion provides a historical overview of approaches to the question of rationality, describes some theoretical positions which result, and presents several frameworks for contemporary investigation. Whereas the primary focus is on the relationship of consciousness, reason, and selfhood from a Western perspective, especially since 17th- and 18th-century Age of Enlightenment, it must be noted that development of the role of reason in Eastern traditions is pertinent but not developed, given space constraints.

Understanding Rationality

There is no single theory of rationality, and over the thousands of years humans have reasoned about the nature of the mind, thought, and cognition, a massive body of philosophical, scientific, metaphysical, and psychological models have been generated. Mario Bunge, noted philosopher-scientist and author of the eight-volume *Treatise on Philosophy*, finds that the lack of consensus as to what rationality is and what it can be expected to accomplish is due to the fact that the word describes at least seven different concepts.

These are as follows:

1. Conceptual (aiming for precision and measurability),
2. Logical (aiming for consistency),
3. Methodological (working to question, prove and justify),
4. Epistemological (providing empirical, factual support that is consistent with other facts),
5. Ontological (providing a worldview consistent with existing knowledge),
6. Valuational (aiming for goals that both could and should be attained), and
7. Practical (being instrumental in attaining those goals).

Any one of these definitions or conceptual frames taken in isolation from the others only partially describes rationality, with the consequence that applications based upon those understandings are what Burge labels as only *pseudorational*.

Through the 20th century, essentially metaphysical and supernatural phenomena have often been excluded in consideration of what is strictly *rational*, relegated to realms of speculation and imagination that include mysticism, religion, ethics, superstition, intuition, and similar states of mind. However, as will be seen, recent work in neuroscience, computer science, critical theory, and similar disciplines are blurring those boundaries to some extent.

Rationality Before the Common Era

Philosophers have long declared that human beings are inherently rational. Aristotle, in the 4th century B.C.E., defined human beings as the *thinking animal*, and in 1660, Blaise Pascal called humans *thinking reeds*. Even so, in prehistory and to some extent in classical Greek philosophy, the universe itself was the source of all intelligence—observing, controlling, and acting through the material world, including human beings. Decision-making was the corollary of divination. Humans searched in bones, entrails, smoke, and dreams for signs and portents and sought direction through magical divination or instruction passed from the First Cause or unmoved Mover through prophets, priests and priestesses, and oracles.

Evidence of human rationality greatly predates theorizing about it, with written evidence beginning around 3500 B.C.E. The inquiry into the nature, scope, and acquisition of knowledge became a distinct philosophical genre, *epistemology*, also dealing with related concepts such as perception, memory, belief, and justification. Thales of Miletus (c. 620 B.C.E.–546 B.C.E.), often referred to as the first philosopher, believed that the earth floats on water and that “all things are full of God,” and is

credited as being the first to assert that the human soul is immortal.

Pythagoras (c. 570 B.C.E.–c. 495 B.C.E.) saw the natural world as being ordered according to a set of mathematical ratios. He promoted a way of life that included a strict moral code, which included the mandate to train the memory, and his teachings influenced both Plato and Aristotle. Heraclitus (535 B.C.E.–475 B.C.E.) postulated universal *logos* (pure rationality), which humans could tap into but which most chose to ignore and to continue as sleepwalkers. Epicurus (341 B.C.E.–270 B.C.E.) questioned the usefulness of formal logic and developed a hedonistic approach to life that involved the use of reason primarily to avoid fear. Plotinus (204 C.E.–270 C.E.) was a Neoplatonist who believed that the human mind was primarily rooted in the material world but was nevertheless capable of experiencing *the One*.

Even now, there is a question as to whether rationality can be taught and, if so, how would such progress in critical thinking be effected and measured. Two paramount philosopher-teachers of rationality were Socrates (c. 470 B.C.E.–c. 399 B.C.E.) and his disciple Plato (c. 428 B.C.E.–c. 347 B.C.E.), who presented Socrates' ideas in the form of *Socratic dialogues*. Two key elements of Socrates' philosophy of rationality presented in Plato's dialogues are the *Theory of Forms* and the *Allegory of the Cave*. For Plato, knowledge or truth separate from simple opinion or belief exist but is predicated upon absolute, unchanging pre-existing Forms. What we glean from our experience, thought, and reflection has been called into material being as reflections of those ideal realities. Since the eternal Forms are antecedent, then our immortal souls had full knowledge of them but *forgot* that absolute ideal world at birth. Thus, *learning* is a process of simply remembering.

Plato's *Allegory of the Cave* is a metaphor for existence as a state of coming to know what the soul already knew. Here, Plato describes humans as prisoners chained in a cave facing a wall of shadows which they believe to be reality. They don't understand that the shadows they see are mere reflections of real things being moved around behind the prisoners outside of their field of vision but in front of a fire that is casting the shadows.

The prisoners can break free (i.e., the mind can escape the Cave) only by coming to understand the Forms which constitute true reality that prevails outside the Cave in sunlight.

Like Plato and Socrates, Aristotle (384 B.C.E.–322 B.C.E.) believed that the greatest attainment of human life was possible only in attempting to acquire knowledge through reason or *logos*. According to Aristotle, all of natural creation is arranged in hierarchically ordered taxonomy or "Great Chain of Being," with humans at the crown based upon our complex intellectual abilities. He explains the relationship of body and soul (*psyche*) with the analogy that if the eye were an animal, sight would be its soul. Reason, or *nous*, is the highest form of rationality; the *unmoved Mover* of creation is divine *nous*. Aristotle conceived of two forms of human knowledge, *sophia* or wisdom and *phronesis* or practical knowledge; these correspond generally to conceptions of theoretical and instrumental rationality. In stating that *every primary substance falls under a secondary substance*, Aristotelian logic foreshadowed today's concepts of epistemological *supervenience* (discussed subsequently).

Concurrent with the emergence of Western epistemological models of human reason, a set of teachings that may be termed *dharmic philosophies* developed in Asian cultures. These, central to Buddhism, Jainism, and Hinduism, advocate practices (such as meditation) that develop *mindfulness* as in Buddhism and, with the spread of these teachings to China, encourage efforts to manifest Confucian *reasonableness* as opposed to logical rationality. In the 6th century B.C.E., the Chinese sage Laozi (in the older, Wade-Giles transliteration, Lao-tzu) taught his followers to conform to events in their lives through non-willful action, passively but consciously following the Way (the Dao).

Western Early Modern Philosophers on Rationality

Francis Bacon (1561–1626), English statesman, scientist, and philosopher, emphasized *inductive reasoning* (building a body of knowledge through empirical sense experience in the natural world),

furthering development of the scientific method as well as establishing the method of inductive logic. He said *Knowledge and human power are synonymous* and believed humans had two souls, one animal and one rational. Bacon viewed the intellect to be both lazy and restless by nature, subject to committing errors in reasoning under the influence of four types of mental *idols*: idols of the tribe, the marketplace, the cave, and the theater. Bacon strongly influenced such eminent philosophers as John Locke, David Hume, John Stuart Mill (1806–1873), and Bertrand Russell (1872–1970).

René Descartes (1596–1650), the *father of modern philosophy*, established the school of thought now known as *rationalism*, postulating that the ultimate source of knowledge is reason. This was a move away from faith in classical authorities and toward reliance upon the individual's powers of reason, especially deduction. He believed some truths exist a priori and can be known innately through the power of reasoning, verified through sense knowledge, experience, and experimentation.

In *Discourse on Method*, Descartes describes a formula to arrive at these conclusions using four rules. The first was “never to accept anything as true that [he] did not know evidently to be so.” The second was to take a mathematical approach and divide that which he was examining “into as many parts as possible,” dealing with each individually. The third was to order his thoughts from the simplest and easiest to know gradually to the more complex. And the last was to make “enumerations so complete and reviews so general that [he] would be sure of having omitted nothing.” After rigorously employing this method for several months and discarding everything he could not strictly confirm and verify, Descartes was left with one truth “so firm and so certain” that he could not shake it: *Cogito, ergo sum* (“I think, therefore I am”).

In the centuries since Descartes arrived at his core truth, other philosophers have concluded that the *ergo* clause does not necessarily follow the *cogito*—that thinking does not necessarily prove *being*. Even so, Descartes' method of *radical doubt* and its division of all that is into the categories *res logicans* (things thought) and *res extensa* (everything else) established the humanistic

approach to rational explanations of reality that undergirds the modern Western tradition.

Historian and political philosopher Thomas Hobbes (1588–1679) was Descartes' contemporary, both conferring with and critiquing him. Hobbes termed his theories of moral philosophy and of rationality as a form of computation. His interest in the workings of the mind through sensations and imaginings, especially as expressed in language, brought him to appreciate the importance of linguistic *signification*, or *naming*, to instrumental rationality. His conceptualization of the human as a *thinking body* was central to his theory of *materialism*; Hobbes did not believe that a thought could exist without a body to think it. His major work, *Leviathan*, elaborated his conception of a *social contract theory* which, similarly to the Greek *polis* of Aristotle, requires the cooperation of free, equal, and rational participants.

In writings opposing Hobbes's conservative philosophy, the politically idealistic John Locke established (British) *empiricism*. He asserted that all knowledge results solely from the individual's experience of the world apprehended through the senses; thus, he advocated experimental science as an appropriate use of reason. The mind, according to Locke, was “an empty cupboard,” a *tabula rasa* or blank slate, until the senses imprinted experience on it. Locke believed that some aspects of existence could be apprehended through reason, while others could not; his treatise *Essay on Human Understanding* addresses this dichotomy.

“Heretic extraordinaire” Baruch de Spinoza (1632–1677) developed a radical philosophy that included elements of Stoicism, Talmudic principles, and Cartesian duality. He was attacked and excommunicated from his Sephardic Jewish community for denying the existence of the soul or of an Old Testament deity (“a God existing only in the philosophical sense”) or of a body of received law or morality binding on humans. However, his theory of rationality—that passion and emotion should be overcome so that all things might be deduced by the rational mind from a few basic principles—could be considered more *pantheistic* than atheistic, in that he believed all material existence was imbued with and part of a God-force. He used foundational axioms and definitions to

differentiate between two forms of knowledge: *ratio* (rational) and *scientia intuitiva* (intuitive).

Scottish philosopher and historian David Hume (1711–1776) is noted for study of what has become known as *philosophical naturalism* and *moral rationalism*. He rejected Descartes' notion of a priori knowledge and believed that what we can know through the mind will only be known as a result of the impressions of the world processed through experiences. His work influenced scholars such as economist Adam Smith, biologist Charles Darwin, and Utilitarian philosopher Jeremy Bentham.

Gottfried Leibnitz (1646–1716) was a polymath and diplomat whose love of mathematics, especially geometry, eventually led him to develop novel systems of calculus. Subsequent to conversations with Spinoza in 1676, he turned his attention more fully to philosophy. Perhaps his most important contribution to philosophy of mind is his notion that preestablished harmony between mind and body is the basis of rationality. His reflections constitute a critique of both Cartesian dualism and of rationalism. By including consideration of human striving, motivation entered epistemological philosophy.

The German philosopher Immanuel Kant (1724–1804) is a central figure in philosophy generally and especially in consideration of the influence his *Critiques of Pure Reason* (the original leaning toward metaphysical idealism and the revision more toward realism) have had on thinkers who have followed him. He attempted to delineate the sources and conceptual categories of human knowledge and to elucidate a method for acquiring and applying it. Kant rejected theories asserting that experience was the necessary foundation for rational thought insofar as he saw certain categories such as theology, psychology, and cosmology as existing but being beyond the realm of human sense experience.

Georg Wilhelm Friedrich Hegel's (1770–1831) early studies focused on both philosophy and theology, with the result that the scope of his writings spans and integrates multiple domains. Using logic as a starting point, Hegel searched for a philosophical system which could attain *absolute knowledge* that would answer all questions—epistemological, ontological, and metaphysical. The Hegelian system conceives of all things,

including the human mind and its capacity for rationality, as embedded within pure, conscious, eternally active Spirit or *Geist*.

Postmodern Rationality

In the modern-to-postmodern paradigm shift in ways of thinking about thinking and what it means to be *rational*, previously established *givens* are increasingly contested. Although most would not disagree that humans have an innate ability to perceive situations, alternatives, and potential outcomes as well to ruminate upon a broad range of immaterial abstractions through cognition and reason, arriving at a common *truth* as a result of this philosophizing is no longer expected. The consensus currently is that attempting to delineate a set of universalized *norms* for rationality will not succeed, for a number of reasons. For example, there are disputes about how rationality should be defined, what rationality should do, and how rationality should be measured or judged based upon consensually defined norms. And with respect to the latter, philosopher James Broome reasons that *rationality* may not be *normative* at all. As a result, a variety of *isms* tied to rationality has developed, a few of which are described below.

Foundationalism

Proper or well-founded beliefs are necessary ingredients for effective reasoning. Some beliefs are *basic* and provide justification for further elaboration. Aristotle's *Posterior Analytics* and Descartes' *cogito* rest upon a foundationalist approach. Rationalism and empiricism are not at odds with this model. A philosophical debate as to the necessary interconnection of these basic beliefs was generated in the mid-20th century, metaphorically arguing, for example, whether a raft of beliefs or a web of beliefs would produce the most justifiable derivative (or secondary) beliefs.

Perspectivalism (or Perspectivism)

As enunciated generally by Friedrich Nietzsche and further developed by John Frame, Vern Poythress, and Jonathan Kvanvig, this model lies somewhere between *relativism* and *absolutism*. A corollary to *foundationalism*, it states that it is impossible for an individual to be objective, since

each person experiences the world from a unique subject position. Thus, many ways of seeing and reasoning exist, some more valid than others, but none is objectively true. Nietzsche (1844–1900) says, “Insofar as the word ‘knowledge’ has any meaning, the world is knowable; but it is interpretable otherwise, it has no meaning behind it, but countless meanings. . . . There are many sorts of eyes. The sphinx too has eyes; consequently, there are many sorts of ‘truths,’ and consequently there is no truth” (*Will to Power*, Section 540).

Coherentism

The *coherence theory of justification* (which differs from the *coherence theory of truth*) requires that various underlying beliefs arise out of a body of beliefs consistent with each other. This thesis is generally considered to be at odds with *foundationalism*, generating a variety of arguments for and against. Some consensus has developed around its relevance to moral epistemology based in part upon interpretations of the work of John Rawls (1921–2002).

Internalism and Externalism

One step in rationalization is to apprehend and justify certain beliefs. In epistemology, the debate framed as *internalism* versus *externalism* has to do with whether the individual can rightly justify a belief based upon mental content is accessible solely from within that person’s consciousness or whether it is necessary that the individual draw upon external sources. Externalists aver that information drawn from outside the interior psychological states of the knower is a necessary condition for knowledge. Both *internalist* and *externalist* models of belief justification include semantic, motivational, and historiographical theoretical subcategories.

Evidentialism

Related to rationality, evidentialism is a type of *supervenience theory*—that is, sets of properties are hierarchically ordered to determine which properties are requisite (which *supervene*), stating that a change in one set of properties will cause a change in related property sets. Evidentialism requires that a belief (or disbelief or withholding

of belief altogether) must be justified by *acceptable* evidence and *good reasons* which support the belief. The *evidentialist* would say that *One ought to believe only what one can believe*. However, this can trigger heated discussion of what constitutes valid evidence and which categories of evidence properly supervene upon other categories. Forms of this position can be seen, for example, in the writings of David Hume and Bertrand Russell.

The Rationality (and Irrationality) of Groups

In a shift from a focus on the rationality of individuals, scholars in the 19th and 20th centuries began developing theories of collective or group rationality. In some cases, though, what emerged were theories demonstrating groups often behave more irrationally than individuals acting alone would be likely to behave.

German philosopher-political economist Max Weber (1864–1920), taking a sociological viewpoint to investigate patterns and processes of *rationalization*, identified four types of rationality (practical, theoretical, substantive, and formal) conforming to his four modes of social action (affectual, traditional, value-rational, and means-end). He conceived of these as universal in some sense but developing in certain ways specific to Western cultures or *life-spheres*. Weber believed that individuals have a genetic predisposition to rational choices based upon self-interest, but that this is acted upon throughout life by a variety of forces, including the mores specific to social situations and cultures, learned values, and *nonrational* factors unique to each individual.

The critical social theory of Jürgen Habermas (b. 1929) addresses a multitude of philosophical, metaphysical, and sociopolitical topics, many based upon his conception of the *public sphere*. His approach integrates philosophical and social scientific methods. Habermas’s conception of rationality is essentially teleological, linking reason to fulfillment of innate urges and needs. Dialectical engagement between individuals and society by way of *communicative action* is oriented toward fulfillment of collective needs. Habermas describes *rationality* as an evolutionary learning process by a society, the product of interpersonal discourse rather than solitary reasoning processes. His

interest lies less in how *knowledge* is constituted than in “how speaking and acting subjects acquire and use knowledge.” His *discourse theory* depicts application of rationality to collective learning.

Some findings of social psychologists studying groups find that the tendency toward conformity in groups illustrates human *irrationality* rather than rationality. An example is Irving Janis’s (1918–1990) concept of *groupthink* (first published in 1972), illustrated by such events as the Bay of Pigs invasion of Cuba in the 1960s and, more recently, decisions related to Iraq and its former leader Saddam Hussein. Groupthink demonstrates that highly cohesive groups are prone to abandon rational decision-making processes to maintain unity. Studies of groupthink evidence the positive role conflict, properly managed, can perform in promoting group reasoning and rationality. Classic studies of group behavior such as Solomon Asch’s 1953 line-length experiment, Stanley Milgram’s 1961 obedience-to-authority studies, and Philip Zimbardo’s 1971 Stanford prison experiments demonstrate a disturbing tendency for members of groups to abandon logical and moral reasoning to follow the crowd or an authority figure. Research in this area is ongoing.

Recent Perspectives on Rationality

Many organizations are interested in models of practical rationality that can further critical and near-constant decision-making processes. In *real-world* application, constraints such as limited resources (especially time to arrive at a conclusion) and, in small-group decision-making, the necessity for individuals to cooperate selection and implementation, that purely rational, logical operations are not particularly effective. A quasi-logical model, *bounded rationality*, first theorized by James March and Herbert Simon in 1958, offers *satisficing* as an alternative to the *optimizing* approach, which searches for the ideal solution. In satisficing, a *good enough* solution is the goal. Simon observes that intuitive processes and experience-based knowledge—which are neither logical nor illogical—are valid elements of workable decisions.

For the French historian, philosopher, and social critic Michel Foucault (1926–1984), the philosophical system that attains the power to determine what *knowledge* is becomes embedded in the philological

system (structural elements of language or *langue*). In this way, episteme acquires the power to constrain and bound what is right or good to think or know. This hobbles the evolution and vitality of rational understanding of the “lived experience of an individual.” Foucault’s notions of *bio-power* and *bio-politics*—wherein power systems that build upon established norms and received knowledge take control over bodies and entire populations. He extends this critique to those who allow themselves to be physically and mentally *disciplined* in this way.

The emphasis in Western cultures on the supremacy of logic also has been subjected to critique by feminist and critical-cultural theoreticians. Dennis Mumby and Linda Putnam take the standpoint of *critical organizational theory*, accepting elements of Herbert Simon’s (1916–2001) theory of bounded rationality as a valuable critique of *pure rationality*, but still finding it problematic because its core assumptions are rooted in the *male-centered patriarchal worldview* that overprivileges the rational, relegating the emotional or intuitive to inferior status. They call for accepting *bounded emotionality* as both a corollary and an alternative to bounded rationality; they urge a *rewriting* of the relationship between reason and emotion.

Scientific and technological advances in neuroscience and computer science call into question some old ideas about rationality and emotion prevalent in the Western tradition. Neurobiologist Antonio Damasio (b. 1944) foregrounds ongoing laboratory research into *affective neuroscience* and sees feelings as having a truly privileged status because they are emanations of the sense experience of the body. Thus, emotions are the *causal link* among events, experiential states, and rational thought.

Augmenting *truth* with *truthiness* (the Merriam-Webster word of the year for 2006, meaning “what one wishes to be true rather than what one can prove to be true”) is recognized as a practical 21st-century discursive strategy. Concurrently, some paths that began to diverge circa the 6th century B.C.E. appear now to be moving toward convergence. Scholarly and scientific studies increasingly find compatibility between Eastern and Western views, especially in terms of *phronesis*, or practical wisdom. Together with some acceptance of the emotional and intuitive as valid ways of *knowing*, experimental and theoretical studies find mindfulness and reasonableness to be forms of perception and sources of wisdom at least

as effective as focused reason and logical processes. Weick and Putnam outline some of these claims in their (2006) article “Organizing for mindfulness: Eastern wisdom and Western knowledge.”

Leslie J. Reynard

See also Abstract Knowledge; Analysis; Concept; Deduction; Epistemology; Induction; Reasoning; Speculation; Understanding

Further Readings

- Bermudez, J. L., & Alan Millar, A. (Eds.). (2002). *Reason and nature: Essays in the theory of rationality*. Oxford University Press.
- Broome, J. (2007). Is rationality normative? *Disputatio*, 2(23), 161–178.
- Bunge, M. (2013). Seven desiderata for rationality. In J. Agassi, & I. C. Jarvie (Eds.), *Rationality: The critical view* (Nijhoff International Philosophy Series, Book 23). Springer, 5–15. (Original work published 1987)
- Cialdini, R. B., & Goldstein, N. J. (2004). Social influence: Compliance and conformity. *Annual Review of Psychology*, 55, 591–621. <http://dx.doi.org/10.1146/annurev.psych.55.090902.142015>
- Damasio, A. (1994). *Descartes' error: Emotion, reason, and the human brain*. Penguin.
- Foucault, M. (1972). *The archaeology of knowledge and the discourse on language* (A. M. Sheridan Smith, Trans.). Vintage Books.
- Habermas, J. (1984). *The theory of communicative action, Vol. I: Reason and the rationalization of society* (T. McCarthy, Trans.). Beacon.
- Kalberg, S. (1980). Max Weber's types of rationality: Cornerstones for the analysis of rationalization processes in history. *American Journal of Sociology*, 85(5), 1145–1179.
- Kim, J. (2010). *Philosophy of mind* (3rd ed.). Westview Press.
- Kvanvig, J. L. (2014). *Rationality and reflection: How to think about what to think*. Oxford University Press.
- Mumby, D. K., & Putnam, L. L. (1992). The politics of emotion: A feminist reading of bounded rationality. *Academy of Management Review*, 17(3), 465–486.
- Simon, H. A. (1972). Theories of bounded rationality. In C. B. McGuire & R. Radner (Eds.), *Decision and organization*. North-Holland.
- Weick, K. E., & Putnam, T. (2006). Organizing for mindfulness: Eastern wisdom and Western knowledge. *Journal of Management Inquiry*, 15(3), 275–287.
- Woelfel, J., & Trulove, S. (1991). *Patterns in western civilization*. Ginn.

REALISM IN MATHEMATICS

The most general, least controversial, and consequently also least informative description of the doctrine of mathematical realism is the view on which mathematical objects (such as numbers, sets, right triangles, and tensors) are *real* (in the sense of not made up, either deliberately or unconsciously, by human subjects). To get a sense of why the doctrine is interesting enough to merit an entry here and substantive enough to remain the topic of ongoing debate among mathematicians and philosophers of mathematics, it is helpful to explore some of its more specific, and thus more controversial versions and their consequences, and to consider implications of the discussion of mathematical realism for realism *simpliciter*.

Mathematical Platonism

For better or worse, mathematical realism has long been associated with the classical Athenian philosopher Plato (427–346 B.C.E.), who in his dialogue *Meno* famously has his mentor Socrates lead an untutored slave boy through a series of basic geometric proofs. Since Socrates claims simply to have asked questions with a view toward soliciting what his interlocutor already knew, without ever instructing him in any way, he concludes that for all his lack of education the boy already knew some mathematics. Thus, by extension, we all have access to mathematical knowledge, though such access may remain merely latent. Although this argument appears to be about *mathematical knowledge*, it is also about *mathematical objects*, for it is in virtue of the actual mathematical objects to which the slave boy gains access that the insights he reaches are true and thus count as knowledge. The mutual dependence of accounts of mathematical knowledge with accounts of the nature of mathematical objects can give pause even to those philosophers inclined in general to treat metaphysics and epistemology as independent enquiries. We will return to this point later.

The metaphysical approach to mathematical objects associated with Plato's *Meno* is sometimes called mathematical *Platonism* to distinguish it from other forms of mathematical realism. For the

Platonist, mathematical objects are real, but of a very different order from other sorts of things commonly claimed to be real, including medium-sized physical objects, electrons, and nation-states. The Platonist takes mathematical objects to be eternal and unchanging, perhaps lacking even the property of being one day older today than they were yesterday. Platonism has the advantage of being consonant with the apparent universality of mathematical truths, including their invariance with respect to place and time. Among its disadvantages is the difficulty of explaining how creatures that are far from eternal and unchanging, such as human beings, could interact with mathematical objects in a way even remotely analogous to the ways in which we interact with other objects of knowledge. To see medium-sized physical objects, for example, we perceive the photons that bounce off their surfaces, enter our eyes, and onto our retinæ. Whatever faculty allows us to gain knowledge of eternal, unchanging mathematical objects, it would have to be very different from sight.

In modern philosophy of mathematics, one of the more influential mathematical Platonists is German mathematician, logician, and philosopher Gottlob Frege (1848–1925), who along with his contemporary Edmund Husserl (1859–1938) helped restore mathematical realism to prominence by critiquing the prevailing mathematical psychologism (on which mathematical objects are contingent aspects of human psychology) of the preceding generation. In an influential essay, Frege consigned mathematical objects to a *third realm*. Mathematical Platonists like Frege believe mathematical objects are real, and that they subsist neither in the realm of physical objects, nor in the realm of *mere* ideas, constituents of human psychology. They belong to a third realm of abstract objects.

Realism and Idealism

Since the 17th century, in metaphysics as well as in the philosophy of science, it has been customary to distinguish *realism* from *idealism*, treating the two doctrines as diametrically opposed. Some discussions even presuppose an excluded middle between the two doctrines, such that between them they must encompass the entire range of possible positions. In general, *realism* is the view that many ordinary objects of human cognition

are real, independent of human thought. In general, *idealism* is the view that all there is in the world are minds and their ideas. In the philosophy of science, *realism about x* is the view that x is real and independent of human thought, where for x we may insert any object of scientific knowledge or enquiry. In the philosophy of physics, for example, we find proponents of realism about electrons (and other particles resistant to direct observation), fields, and space-time. In the philosophy of biology, there are those who argue for realism about species and ecosystems. By contrast, *idealism about x* is the view that x is an aspect (or constituent, or function) of minds and their ideas, rather than of any mind-independent world. Such idealisms come in many different flavors. For a Berkeleyan idealist (after the British philosopher George Berkeley, 1685–1753), to be is to be perceived. Perception itself constitutes the perceived objects, which have no reality independent of their status as objects of perception. By contrast, while remaining agnostic about the existence of things-in-themselves (things as they are independent of our perception or knowledge of them), a post-Kantian, transcendental idealist (cf. the Prussian philosopher Immanuel Kant [1724–1804]) about space and time might assert that space-time, rather than an unfolding and expanding constituent of a cosmos far older than and largely indifferent to minds, is instead a necessary constituent of any finite intellect capable of experiencing any part of the cosmos in the first place.

Let us turn to *realism* and *idealism about mathematical objects*, say numbers. A realist about the natural numbers would presumably claim that they exist whether or not there is anyone in the universe who uses them for counting. By contrast, an idealist about the natural numbers would assert that the natural numbers are something like ideas in the minds of creatures that count. As this example helps illustrate, the distinction between realism and idealism in mathematics gives us reason to doubt that realism and idealism are anything like polar opposites. If minds and their ideas exist—if they are real—then the idealist about natural numbers is not denying their reality *per se*. Viewed a certain way, Frege's third-realm Platonism is both a kind of realism and a kind of idealism: Numbers are real, *and* they are ideas. Whereas minds may entertain and manipulate such ideas, the existence

of the ideas themselves is not contingent on the existence of the minds engaging with them. Again, mathematical ideas belong not to the realm of mind-independent physical objects, nor to the intracranial or social realm of merely human ideation and communication, but to the (third) realm of immutable abstract objects.

In light of such considerations, cautious discussions of the merits of realism in mathematics are better-off setting out from the distinction between *realism* and *antirealism*, rather than realism and idealism. Some brands of mathematical Platonism may count as varieties of both realism and idealism. Furthermore, there are important antirealist platforms that may well fail to count as varieties of idealism. These include mathematical *nominalism*, *fictionalism*, *instrumentalism*, and *conventionalism*, along with positions combining elements of two or more antirealist approaches.

Benacerraf's Problem

As noted earlier, the domain of mathematics gives pause even to those philosophers of science otherwise inclined to treat metaphysics and epistemology as independent branches of enquiry. In this respect the philosophy of mathematics has much to teach the philosophy of science. An account of mathematical objects that leaves them unknowable, strangling mathematical knowledge in the crib, would be unacceptable; the whole point of such an account is to help explain mathematical knowledge. Conversely, an account of mathematical knowledge that made the objects of mathematical knowledge into something too strange or impossible would be suspect by virtue of its incompatibility with any answer to the question, what is it that mathematicians have knowledge of? At least in mathematics, there appear to be checks and balances between epistemology and metaphysics, a prospect that philosophers of science studying other scientific disciplines might do well to consider.

In his famous 1973 paper, "Mathematical Truth," Paul Benacerraf articulated the tension between reasonable if seemingly conflicting demands on any philosophical account of mathematical knowledge in a way that continues to shape this conversation. To count as knowledge, mathematical beliefs must be (at least) true and

justified. Let an account of the truth of sentences count as a stand-in for an account of the truth of beliefs. On this assumption, Benacerraf claims, the best available theories of truth are variations on the *correspondence theory*, on which sentences are true just in case the objects corresponding to their nouns are in fact arranged in ways corresponding to their predicates. But by 1973, while continuing to accept truth and justification as individually necessary conditions for knowledge, analytic philosophers no longer held them to be jointly sufficient. Thanks to the American philosopher Edmund Gettier, beliefs worthy of being called *knowledge* now had to be not only true and justified, but also *caused in the right sort of way by the objects of belief*. This further stipulation is sometimes called the *causal condition* for knowledge. If truth in general is correspondence, then the truth of mathematical theorems and beliefs should properly be understood as a function of correspondence with mathematical objects. Such correspondence, in turn, presupposes *realism about mathematical objects*: If the objects weren't real, they couldn't correspond with anything in language, or anything in the minds of knowing subjects.

Mathematical theorems are not, however, merely contingently true, in the way it is contingently true, as I write this, that I am wearing a blue T-shirt. They appear to be *necessarily* true; that is, invariant with respect to time and place. So if they are true by virtue of correspondence with mathematical objects, *where are those objects*? Perhaps they reside in Frege's third realm, a realm of nonspatiotemporal abstract entities. If so, Benacerraf argues, this raises a further problem, for on the best available epistemologies, the causal condition is a necessary condition for knowledge. Satisfying the causal condition, however, requires the interaction between the epistemic subject and the object of knowledge. With *ordinary* objects of knowledge, such interaction seems unproblematic: I interact with my car by visually perceiving the light bouncing off its rust-pocked surface, hearing its misfires, or smelling its burning engine oil. By way of my senses, my car *causes* me to have beliefs about it, such as my belief that it is time for it to stop being a car and become scrap. How, though, do I imagine I might causally interact with third-realm entities? I can see three oranges in the usual way, by photons that bounce off their surfaces,

and I can likewise see three apples. However, the number three itself, the property of threeness, or the pure set of cardinality three, do not seem to be visible to me, or otherwise accessible to my ordinary senses.

The Austrian mathematician Kurt Gödel (1906–1978) posited that our access to mathematical objects must be mediated by a faculty he called *mathematical intuition*, like perception in some ways, though not dependent on sensation. Leaving aside the questions of how such a faculty might arise or what would make it reliable, there remains the fundamental problem of how spatiotemporally bounded creatures such as ourselves could ever gain access to the nonspatiotemporal residents of the third realm. More recently, the American philosopher Penelope Maddy has argued that the role Gödel envisioned for mathematical intuition could be played by ordinary perception. Consider visual perception. We claim to see cars, meerkats, and broccoli stalks by virtue of our ability to perceive photons that bounce off their surfaces onto our retinæ. In such interactions we only ever directly engage with the proximal surfaces of the objects in question and never with the whole objects themselves. If we are nonetheless prepared to allow that we perceive those objects, Maddy argues, then there is no good reason for denying that we perceive (certain) mathematical objects. Maddy's approach to mathematical realism addresses Benacerraf's problem by moving mathematical objects from the third realm squarely into the first, where reside other objects of perception.

Indispensability

Consigning mathematical objects to the first realm would not by itself force them into the same category as medium-sized physical objects. Consider another sort of first-realm denizen about which realism appears perfectly warranted, even if direct observation is impossible. Protozoa are single-celled creatures too small to see with the unaided eye, though they can be seen under sufficient magnification. Electrons, however, are not merely too small to see. According to our best current physics, they are unobservable in principle. Yet we are justified in believing in electrons, because their existence is indispensable to the coherence of the very same current physics that decrees them

unobservable. For its part, such physics does not want for observational support. Physics is an empirical science, founded in evidence. Although that evidence does not include the direct visual inspection of electrons, it does include, for example, the observational records of Robert Millikan's oil-drop experiment, whence the magnitude of the unit charge of the electron was first extrapolated. American philosopher Willard V. O. Quine's 1968 essay "Epistemology Naturalized" is a canonical source for this argument that we are justified in believing in the reality of electrons—or indeed in other observationally challenged entities—because they are indispensable to our current best empirical science. An extension of this argument in support of realism in mathematics seems perfectly natural. If indispensability-in-science militates for the reality of electrons, neutrinos, fields, and the Big Bang, it should perhaps weigh even more heavily in support of the reality of integers, real numbers, complex numbers, sets, right triangles, and finite Abelian groups, all of which play pivotal roles in articulating well-confirmed scientific theories.

A brand of realism founded primarily in arguments from the indispensability of mathematical objects in science (sometimes called the *Quine–Putnam indispensability arguments*) may be classified as *indirect realism*, as distinct from a *direct realism* founded in direct perceptual or intuitive access to the objects of knowledge. Another argument sometimes adduced in support of indirect realism is exemplified by Hungarian American mathematician Eugene Wigner's reflections on the *unreasonable effectiveness* of mathematics in the physical sciences. Math works. It works so astonishingly well, in fact, that its effectiveness would be nothing short of miraculous if mathematics were not somehow *about something physical*—if it did not at some level represent fundamental facts about the physical universe. At first blush, indispensability arguments and unreasonable effectiveness arguments may look like close cousins. However, the unreasonable effectiveness arguments accommodate one perplexing historical observation more easily ignored by advocates of the indispensability arguments. The history of mathematics offers numerous examples of entire branches of mathematics developed decades or centuries in advance of their applicability to any

empirical science. For example, the discovery of non-Euclidean geometry is typically dated to about 1830, when Nicolai Ivanovich Lobachevsky (1792–1856) and Janós Bolyai (1802–1860) published seminal papers. By contrast, physicists did not reach consensus about the non-Euclidean character of space-time until after the promulgation of Albert Einstein's theory of General Relativity in 1915. Though the unreasonable effectiveness argument does not by itself explain how pure mathematics can forge so far ahead of its scientific application, it also offers no objections. By contrast, for an indirect realist relying on indispensability arguments for the reality of mathematical objects, the apparent ability of mathematicians to gain knowledge of mathematical objects well in advance of any (indirect) scientific evidence of their reality is inexplicable.

Proof

Let us conclude this entry with one final persistent problem for realism in mathematics. This problem affects several varieties of mathematical realism, but for ease of exposition we shall focus on indirect realism, as explained in the preceding section. For the indirect realist, an argument for the reality of a given collection of mathematical objects Γ (gamma) might turn on the indispensability of Γ in articulating current scientific theory. However, such arguments rarely if ever carry any weight in pure mathematics, who instead rely on *mathematical proof*. In an earlier example, the plausibility of indispensability arguments for mathematical realism turned on an analogy between physical unobservables such as electrons and mathematical objects such as numbers and sets. Due consideration for the importance of mathematical proof in establishing mathematical knowledge immediately shows the limitations of this analogy. It seems likely that there is no single, universal scientific method. Nonetheless, whatever methods underwrite arguments for the existence of electrons seem unlikely to bear any resemblance to the mathematical proofs whereby one might demonstrate the existence of transfinite cardinals greater than $\aleph_0$. Part of the problem is that mathematical proofs, and the mathematical knowledge they underwrite, have long been thought to be *a priori* (prior to experience or based on reason alone).

Empirical science, almost by definition, is a *posteriori* (grounded in experience). However, even a conclusive argument in support of mathematical empiricism (the view that mathematical knowledge is *a posteriori*) would still have to contend with the methodological disparity between mathematics and the (other) empirical sciences.

Alternatively, we might allow consideration of the reality of mathematical objects to take a backseat to the discussion of the importance of mathematical proof in the growth of mathematical knowledge. One possible outcome of this approach is a very different sort of mathematical realism from any of those considered thus far. Let us say, for example, that my house is very real. It is eminently physical, as witnessed by the fact that my homeowner's policy attaches a value to the possibility that it might burn to the ground. But it is not *natural*. It was built—not by me, to be sure, but by human beings. Understood as a kind of mathematical realism, *mathematical constructivism* treats mathematical objects as the very real but also intentional products of human labor: the labor of mathematical proof.

Conclusion

For all the problems facing realism in mathematics, the view continues to have its share of champions. One factor in its continued prominence is the sense that one ought to trust the scientific experts and defer to their practices. In mathematics, the scientific experts are practicing mathematicians, who are often observed talking about their objects of enquiry, be they imaginary numbers or *n*-dimensional knots, *as if they were real*. If the experts think of them as real, who are we nonexperts to challenge them? In the philosophy of mathematics, mathematical Platonism is often asserted, without evidence, as the nearly unanimous position of professional mathematicians. Even if this claim were supported by an empirical survey of mathematicians, it does not change the fact that mathematical Platonism, and indeed mathematical realism in general, are not *mathematical* doctrines. They are *philosophical* doctrines. Mathematicians may well have philosophical opinions, including opinions about the reality of mathematical objects. Some printers employed by the U.S. Mint's Bureau of Printing and Engraving doubtless have views about

the reality of monetary instruments, too. Their views on ink viscosity as a function of temperature lie within their professional expertise. Their views on whether money is real are philosophical views, and as such they may or may not be interesting or compelling, but they merit no particular deference.

Alex Levine

See also Abduction; Applicability of Mathematics in Science; Category Theory; Epistemology; Scientific Realism; Set Theory

Further Readings

- Benacerraf, P. (1973). Mathematical truth. *Journal of Philosophy*, 70, 661–679.
- Frege, G. (1977). Thoughts: A logical enquiry (P. Geach & R. Stoothoof, Trans.). In P. Geach (Ed.), *Logical investigations*. Blackwell. (Originally published 1918)
- Gettier, E. (1963). Is justified true belief knowledge? *Analysis*, 23, 121–123.
- Maddy, P. (1990). *Realism in mathematics*. Oxford University Press.
- Plato. (1980). *Meno* (G. M. A. Grube, Trans.). Hackett.
- Quine, W. V. O. (1968). Epistemology naturalized. In *Ontological relativity and other essays* (pp. 69–90). Columbia University Press.
- Wigner, E. P. (1960). The unreasonable effectiveness of mathematics in the natural sciences. *Communications in Pure and Applied Mathematics*, 13(1), 1–14.

REALISM IN SCIENCE

See Scientific Realism

REASONING

Reasoning refers to the process of drawing inferences or conclusions from observations, facts, or opinions, often by means of logical rules. These inferences when performed correctly serve to advance knowledge, inform decision making and problem-solving efforts in fulfilling human goals of becoming better consumers of information. Reasoning in argumentation simply provides a rationale behind one choice over competing

choices. However, reasoning is not only for convincing others but also for evaluating arguments, and as such, it also has an epistemic function associated with thinking, cognition, and intellect. This entry briefly considers the historical development in the West of some of reasoning's important formal features, then considers the process of reasoning generally.

Etymologically the word comes from Medieval English (*res(o)un*), *reson*, *reisun* and the Old French word *raison* which are all translations from the original Latin stem *veri* meaning “think or reckon.” Thinking is, in essence, a précis of the general processes of cognitive thought, normally associated with human beings. Reckoning, on the other hand, is a specific cognitive process involving judgment and calculation, usually based around some formal system of measure. As a result, in the origins of the word *reason* is a mixture of two primary meanings at once general and formal. Consequently, it should not be surprising then that the development of meanings attached to the word *reason* have continued to mix both general and formal definitions to this one word, depending on circumstance, intent, and in some cases, convenience.

Reason is, by birth of meaning, also related to the development of the words *rate* and *ratify*. Inferences or conclusions are commonly described as being generated deductively or inductively. Additionally, there are numerous other types of reasoning. Because humans constantly draw inferences to justify beliefs and actions, the process or act of reasoning underlies our intellectual activities.

A History of Reasoning in the West

From its very birth, the word *reasoning* has remained both general and specific; formal and informal. This has not only resulted in the word suffering from a recurring lack of clarity and consistency of meaning, but at the same time this has made the word highly attractive for incorporation into philosophical ideologies.

Antiquity

Although before the time of Plato, very little was written about how we reason, it is inconceivable to think that processes and creations such as

arithmetic, algebra, and the Pythagorean theory dating back as early as 4000 B.C.E. existed without the use of reason. However, most attention to the beginning of process of reasoning is credited to the 5th century B.C.E.

The death of the tyrant Thrasybulus in 338 B.C.E. created a need for citizens in ancient Greece to know how to conduct themselves in court so as to recover property that had been seized by the government during the Peloponnesian War. For a price, the Sophists fulfilled that need. The Sophists were itinerant teachers in Athens and the surrounding area who taught, among other things, rhetoric—the art of persuasive speaking. The early Sophist Protagoras, born around 480 B.C.E., taught that knowledge could only be the result of sense perceptions. Protagoras believed that the only reality that existed was that which was expressed by individuals.

Whereas the Sophists often argued with the purpose of winning a case, Plato demonstrated that the use of argument was a method of discovering truth. Plato saw discussion as a way to clarify and organize thought. He also recognized the importance of imagination. He believed that imagination could replace reason when reason fails to supply the necessary connection between the ideal and the actual.

Although Plato did not write specifically about the exact way to reason effectively, he used the teachings of Socrates to illustrate effective arguments. The general methods used by Socrates were called *dialectic*. Dialectic, or the use of dialogue, is a means of intellectual investigation. In most of Plato's surviving works, he has Socrates as a foil for some popular philosophy, often propounded by a Sophist. His method was to ask leading questions, which ultimately demonstrated the fallacy of the particular belief under investigation. Plato's student Aristotle (384–322 B.C.E.), however, took a different approach to philosophy. Aristotle was very interested in perception. He was also concerned with finding ways of making one's perceptions more trustworthy. Aristotle's greatest accomplishment may have been his extensive conceptual development of categorical logic and the syllogistic argument.

The study of logic progressed very little for almost 1,000 years after Aristotle. Famously, the German philosopher Immanuel Kant (1724–1804)

said that it had not advanced at all since then. The Romans were masters at influencing others using the tools of logic and oratory. The idea of the learned man in the Roman world was the citizen-orator. Argumentation formed the basis of rhetoric. Rhetoric was composed of invention, arrangement, style, memory, and delivery, commonly referred to as the canons of rhetoric. The first two were closely related to argumentation. Even so, these modes of reasoning were principally those Aristotle had codified.

The Middle Ages

During the Middle Ages, or medieval period, the dominant social and educational institution was religion. As a result, rhetoric became associated with preaching. After the destruction in 1258 of Baghdad, the capital of the Abbasid Empire at the time, and the subsequent reconquest of Spain by Christian forces, there was very little advance in the study of reasoning. An exception was the development of supposition theory. Supposition theory embraced those properties of linguistic expressions necessary to explain truth, fallacy, and inference, the three core concepts of logical analysis. Peter Abelard (1079–1142) revived human reason as an effective tool for the discovery of truth. A sharp dialectic practitioner, he effectively confronted the predominant realism of the day and replaced it with his own conceptualist philosophy. Conceptualism rejected the reality of universals. Roger Bacon (1214–1294) was an early pioneer of scientific inquiry, expanding the methods that lie at the heart of modern science—experimentation and induction. Equally significant, Peter Ramus (1515–1572) popularized everyday logic. He also introduced the practice of dividing the canons of rhetoric so that invention and arrangement were associated with philosophy. Thereafter, rhetoricians were less involved with argument and more interested in style (gesture, delivery, figures of speech, etc.).

By the late 16th and early 17th centuries, reason quickly developed into a word implying formal systems of argument and thought processes. Suddenly reason, no longer meant simply thinking and reckoning but was regarded as a formal system of thinking and reckoning. By the mid-17th century, reasoning took on the added meaning as

a formal system of proof of thinking and reckoning. No longer was the act of following a certain system or procedure of thinking and/or reckoning sufficient; proof is also required.

The Age of Reason, or the Enlightenment

The 18th century, often referred to as the Age of Reason or the Enlightenment, witnessed a significant change in the way Westerners viewed themselves, the pursuit of knowledge, and the universe. Previously held concepts of conduct and thought could now be challenged both orally and in writing without fear of one's being denounced by the church as a heretic and punished, sometimes even being publicly executed by fire. This was the beginning of a more open society in which individuals were at liberty to pursue individual happiness and freedom, with the power of reason advocated as the primary source for legitimacy and authority. The dominant concepts of the medieval world were being abandoned both politically and socially. Concurrently, extensive changes were occurring in scientific thought and exploration. New ideas were developed, and people were eager to explore these ideas, freely. Immanuel Kant defined enlightenment as the liberation of man from his self-caused state of minority. Minority was defined as the incapacity to use one's knowledge or insight without the direction of another. Reason, philosophically, became regarded as the ability to form and operate upon concepts in abstraction, narrowing information to its basic content, without understanding and explanation.

The Age of Reason saw the introduction of the Scientific Revolution and various progressions of new schools of thought. Dualism, advocated by the French philosopher René Descartes (1596–1650), taught that God (mind) and man (nature) were distinct. Baruch Spinoza, a Dutch philosopher of Sephardic Jewish origin, introduced the idea of *pantheism*—that God and the universe are one—and further that God was a substance consisting of infinite attributes. Believers in Deism, described as the religion of reason, rejected Christianity as a body of revelation, mysterious, and incomprehensible. The Age of Reason gave rise to attacks on basic religious beliefs, rejection of God, and denial of miracles. In an attempt to

dissociate themselves from the mysticism of the Middle Ages, some thinkers during the Age of Reason applauded intellect and scorned spirit. God was believed to be unknowable, if He existed at all, and there was no need for divine communication or revelation. Nature was revelation enough, showing all that needed to be known about God. Human beings were now free to postulate their own theories of existence.

Sir Frances Bacon (1561–1626) rejected Aristotelean logic and Scholasticism in favor of a new framework for discovery, published in his *Novum Organum* in 1620. The new approach consisted of the careful conservation and collection of data, their correct interpretation, and experimentation. *Port-Royal Logic* was another important textbook of logic, published anonymously in 1662. It was written in simple language, contained many new notions, and was highly popular. The French mathematician Blaise Pascal (1623–1662) is presumed to have contributed to the Port-Royal textbook. In collaboration with Pierre de Fermat (1601–1665), he also developed probability theory, which later gave birth to statistics.

Since the time of Descartes, logic has meant formal, symbolic logic. Descartes had an enormous impact on logic. He instituted the idea of systematic doubt. He searched for the basis of reasoning—that which could be accepted axiomatically, without doubt. He found one idea that met that criterion: *I think, therefore I am* (in Latin, *cogito ergo sum*), which he wrote in French as *je pense, donc je suis*). Descartes inserted one other thing into logic—mathematics. He reasoned that God exists from a consideration of Euclid's exposition of geometry. He required that reasoning must proceed from unassailable axioms to conclusions. And then he translated position and motion into mathematical entities.

These ideas were to remain dominant in the sciences and in logic for the next 300 years. Reasoning became identified with formal logic. People began looking for ways to translate informal reasoning into formal logic and, thereby, to infuse reasoning with the certainty of formal thought. Science began looking for axiomatic bases for their investigations. Argumentation became the demonstration of self-evident proof. Other significant thinkers during this period

were Thomas Hobbes (1588–1679), a contemporary of Descartes and the founder of scientific materialism and ethical hedonism; John Locke (1632–1704), who believed that ideas are the products of experience and reflection on those experiences, in other words, reason; and David Hume (1711–1776). Hume proposed that science relies on facts; that things happen in an orderly sequence of causes and effects; that nature is uniform and things are associated naturally on the basis of similarity, proximity, contrast, and temporal association.

The 19th Century

By the 19th century, the word *reason* had become associated with, and entrapped by, the formal system of argument and proof called *logic*. Reason could no longer stand on its own; it had been conquered by the forces of logic. Free thought was no longer to be viewed as identical with reason. Reason was to be regarded as a formal system of thought and reckoning associated with logic. Common sense was no longer to be associated with reason, which had been clearly shown not to rely on a similar definition. Notable contributors to this way of thinking included the English polymath John Stuart Mill (1806–1873) who, in an effort to provide a firm basis for scientific reasoning, formalized inductive logic in his book, *A System of Logic* (1843). His work in logic influenced later logicians. The English mathematician George Boole (1815–1864) noticed a similarity between algebra and logic and developed a calculus of reasoning which was published in his *An Investigation of the Laws of Thought*. It later became known as two-valued or Boolean algebra and became the basis for the mathematics and logic used by digital computers.

The 20th Century

The 20th century saw a number of new philosophical schools of thought, including logical positivism, analytic philosophy, phenomenology, existentialism, postmodernism, and poststructuralism. Some of the notable contributors of reasoning during this period included John Venn (1834–1923), who invented the eponymous diagrams that make syllogistic logic so much easier

to analyze. Friedrich Ludwig Gottlob Frege (1848–1925) invented and axiomatized predicate logic. He was a major advocate of logicism, the position that arithmetic is reducible to logic. David Hilbert (1862–1943) laid the groundwork for formalism, the idea that mathematics is a meaningless exercise, a game, using symbols according to formal rules which are agreed to by consensus *a priori*. The developments in formal logic by Kurt Gödel, Alan Turing, Alonzo Church, and others undermined that program but led to important advances in understanding the distinction, and connection, between mathematical and logical inference. The entire discipline of recursion theory, the basis for much of computer science and our general understanding of modern algorithms, is an outgrowth of this work. Additional contributors to our understanding of reasoning about empirical matters include Karl Pearson (1857–1936); Albert Einstein (1879–1955); Karl Popper (1902–1994); C. L. Hamblin (1922–1985); Douglas Walton (1942–); Thomas Kuhn (1922–1996); and Stephen Toulmin (1922–), who held that formal logic is inappropriate as a model for argumentation. He developed a new and more adaptable model for nonformal reasoning. Toulmin also developed a way to diagram arguments that remains widely in use today. In contrast to Toulmin, others who were disenchanted with linguistic formalism attempted to preserve formal aspects of reasoning, but by different means. Toward the end of the 20th century, Jon Barwise (1942–2000), and his many collaborators and those influenced by him, focused on *formal* aspects of visual reasoning and inference.

The Process of Reasoning

The process of reasoning consists of the following course of action: We make observations. From our observations, we establish facts and/or theories. From facts/theories, we draw inferences. From our inferences, we make assumptions. We use our observations, facts/theories, inferences, and assumptions to form opinions. From our opinions, we then create arguments to defend our opinions. We use analyses to critique our opinions as well as other people's observations, facts/theories, inferences, assumptions, opinions, and arguments.

The Dominant Types of Reasoning

Good arguments derive from two major types of reasoning: deductive reasoning and inductive reasoning.

Deductive Reasoning

Deductive reasoning happens when one works from the more general information to the more specific. As it relates to research or argument preparation, sometimes this approach is called the *top-down* approach because researchers start with a very broad range of information and then work their way down to a specific conclusion. Deductive reasoning can also start with an assumed hypothesis or theory. The hypothesis or theory may be well accepted, or it may be somewhat faulty. Nevertheless, for the argument, it is not questioned. Deduction is also used by scientists who take a general scientific law and apply it to a specific case, as they assume that the law is true. Deduction can also be used to test an induction by applying it elsewhere, although in this case, the initial theory is assumed to be true only temporarily. The goal of deductive reasoning is to arrive at a valid chain of reasoning, in which each statement holds up to testing, but it is possible for deductive reasoning to be both valid and unsound.

The Greek philosopher Aristotle, considered the father of deductive reasoning, wrote the following classic example:

All men are mortal.

Socrates is a man.

Therefore, Socrates is mortal.

In Aristotle's example, sometimes referred to as a syllogism, the conclusion or premise of the argument—that all men are mortal and that Socrates is a man—is self-evidently true. Because the premises establish that Socrates is an individual in a group whose members are all mortal, the inevitable conclusion is that Socrates must also be mortal.

Inductive Reasoning

Inductive reasoning works the opposite way and typically only assigns probabilities to its conclusions. It draws conclusions or inferences from

observations in order to make generalizations. For example, our experience so far is that polar bears are white, and we might use induction to infer with high probability that indeed all polar bears are white. Inferences can be done in four stages:

Observation: Collect facts without bias.

Analysis: Classify the facts, identifying patterns of regularity.

Inference: From the patterns, infer generalizations about the relations between the facts.

Confirmation: Testing the inference through further observation.

Inductive reasoning is sometimes referred to as the *bottom-up* approach. Arguments from inductive reasoning can include part to whole where the whole is assumed to be like individual parts; extrapolations where areas beyond the area of study are assumed to be like the studied area; and predictions where the future is assumed to be like the past. Inductive reasoning is commonly seen in the sciences when people want to make sense of a series of observations. Isaac Newton, for example, famously used inductive reasoning to develop a theory of gravity. Using observations, people can develop a theory to explain those observations and seek out disproof of that theory. Another excellent example of this process in action is the discoveries and works of the naturalist and evolutionary biologist Charles Darwin (1809–1882). He made some observations about how the types of finches vary from each other across the Galapagos archipelago. After some thought and reasoning, he saw that these populations were geographically isolated from each other and that the differences between the subspecies varied over distance. He therefore proposed that the finches all shared a common ancestor, and had evolved and adapted, by natural selection to exploit vacant ecological niches. This resulted in evolutionary divergence and the creation of new species, the basis of his famed work *On the Origin of Species*. This was an example of inductive reasoning, as he started with a specific piece of information and expanded it to a broad hypothesis. Science then used deductive reasoning to generate testable hypotheses and test his ideas.

Tacit Versus Explicit Reasoning

Reasoning occurs at two different levels of awareness: tacit and explicit reasoning processes. Tacit processes that facilitate reasoning occur without conscious intervention and outside awareness; they typically do not require attention. Explicit reasoning processes, on the other hand, occurs within the sphere of conscious awareness. We are aware not only of the outcome of our thinking (as with tacit processes) but also of the processes themselves. Such thinking is often described as *strategic*. It typically requires effort and it allows us to bypass the relatively slow accumulation of experiences that underlie tacit learning. Hence, reasoning involves both tacit (unconscious) and explicit (conscious) processes.

Additional Types of Reasoning

Many types of mental activity used to draw conclusions fall under one or the other of those two classes. For example, in analogical reasoning, an analogy for a given thing or situation is found, where the analogy is like the given thing in some way. Other attributes of the analogical situation are then taken to also represent other attributes of the given thing. Our brains work by patterns and association—if a perception fits roughly into an existing pattern, then the existing pattern may be taken as definitive. To use an analogy, start with a target domain where you want to create new understanding, find a general matching domain where some things are similar to the target domain, find specific items from the matching domain, and transfer attributes from the matching domain to the target domain. Thus, A is like B. M is in A. N is in B. So M is like N. And in cause-and-effect reasoning, one appeals to known or suspected causes either to predict or to explain effects. Cause-and-effect reasoning is generally persuasive as it helps answer the question *why* something happens, making a statement objective and rational rather than a blind assertion.

In addition to the varieties of reasoning that conform to either inductive or deductive inference, Charles Sanders Peirce (1839–1914), and before him William Whewell (1794–1866), suggested another distinct class, *abductive inference*. The general idea is to infer a complete pattern from isolated instances of the pattern. In that way, it is similar to

inductive reasoning. The difference can be illustrated by an example: It was known through the exercise of inductive reasoning where various planets would be observed to be at various times, using Ptolemaic or Copernican systems. But Kepler inferred abductively that, in fact, the orbits of the planets were ellipses. This inference to an entire orbital structure was unprecedented and functioned differently from traditional induction.

Janice D. Hamlet

See also Abduction; Deduction; Induction; Rationality

Further Readings

- Amoretti, M. C., & Vassallo, N. (Eds.) (2012). *Reason and rationality*. Ontos Verlag.
- Barnes, H. E. (1965). *An intellectual and cultural history of the western world* (3rd ed.). Dover.
- Broome, J. (2013). *Rationality through reasoning*. Wiley-Blackwell.
- Leighton, J. P. (2004). Defining and describing reason. In J. P. Leighton & R. J. Sternberg (Eds.), *The nature of reasoning* (pp. 3–11). Cambridge University Press.
- Lohman, D. F., & Lakin, J. M. (2011). Reasoning and Intelligence. In R. J. Sternberg & S. B. Kaufman (Eds.), *Cambridge handbook of intelligence* (2nd ed., pp. 419–441). Cambridge University Press.
- Peirce, C. S. (1993). *Reasoning and the logic of things* (K. L. Ketner, Ed.). Harvard University Press.
- Read, S. (2019). Medieval theories: properties of terms., *The Stanford Encyclopedia of Philosophy* (Spring 2019 Edition), Edward N. Zalta (ed.). <https://plato.stanford.edu/archives/spr2019/entries/medieval-terms/>.

RECEIVED VIEW OF THEORIES

See Theories, Syntactic Conception of

REFUTATION

Refutation (R) can be defined as a procedure to confront a theory with evidence proving this theory is wrong or false. R is a concept rooted in an important tradition of thought, namely

falsificationism, according to which the paramount requirement for a theory to be considered as scientific is that the theory must be refutable. A theory that is not refutable by any conceivable event is nonscientific. R is, therefore, a criterion for distinguishing between statements or systems of statements that are scientific versus other statements, such as those that are pseudoscientific, metaphysical, or religious in nature. There is disagreement, however, as to the meaning of what it is for a theory to be confronted by evidence, as well as what should be the criterion of demarcation between science and not-science. This entry presents the competing views on this topic and discusses the links between refutations and other arguments concerning falsificationism, such as conjectures, falsifiability, and verisimilitude.

Procedures to Confront a Theory With Evidence

The meaning of a theory's being confronted by evidence is a contested topic, marked by various competing views. Because there are no very precise boundaries between these views and because the term R is embedded within falsificationism, it is helpful to map an account of these approaches. Moreover, there exists an open-ended discussion as to whether a basic translation between those approaches might be adopted to allow for a temporary suspension of the general problem of incommensurability.

Keeping this in mind, however, it may be claimed that there are several variations of four types of approaches on the sort of procedures used to confront a theory with evidence, namely, *falsificationism*, *verificationism*, *confirmation holism*, and *pragmatism*. These approaches differ in many ways. One of the most important differences between falsificationism, verificationism, confirmation holism, and pragmatism has to do with how each of these perspectives conceives the act of confronting a theory with evidence. Briefly, from falsificationism's viewpoint, to confront a theory with evidence is formulated as testing this theory with certain possible results of observation that are incompatible with that theory. If observations show that the predicted effect is definitely absent, then the theory must be considered as refuted. In a different way, verificationism posits that confronting a theory with evidence involves identifying

propositions that can be verified by true statements based on observation. Therefore, confronting a theory with evidence is defined as the verifiability of this theory in terms of observation statements. The third approach, confirmation holism (also called the *Duhem-Quine thesis*), claims that for any putative falsification, it is always possible to preserve a scientific hypothesis by revising auxiliary hypotheses in its stead. Finally, pragmatism calls for practical decisions and practical arguments rather than for theoretical reasons or evidence to refuse or accept a theory.

These differences concerning what it means to confront a theory with evidence are related to other differences among these views. One such difference lies in their lines or criterion of demarcation between scientific statements and other sorts of statements—for instance, philosophical. From a different criterion of demarcation to R, that is, verification, Ludwig Wittgenstein tried to show that all so-called *philosophical propositions* were actually *nonpropositions* or *pseudopropositions*, in that they were senseless or meaningless. This idea involved a characterization of science as opposed to philosophy. According to Wittgenstein, all genuinely meaningful propositions must be fully reducible to *elementary* or *atomic* propositions; these are simple statements that describe possible states of affairs and can, in principle, be established or rejected by observation. To confront a theory with evidence means that the theory can be verified by statements based on observation. This implies that the propositions that legitimately belong to science are deducible from true observation statements; statements that may possibly fall within the province of science can possibly be verified by observation statements. These statements coincide with the class of all genuine or meaningful statements. According to this approach, therefore, verifiability, meaningfulness, and scientific character all coincide.

Wittgenstein's verifiability criterion was intended to play the part of a criterion of demarcation. This criterion of demarcation was verifiability or deducibility from observation statements. In this sense, verificationism is a rival view to falsificationism: Falsificationism looks for observation statements that can override the initial hypothesis (and if they are not found, the hypothesis will have greater verisimilitude and, therefore, it will be closer to the truth). By contrast, in

verificationism, it is considered that observation-based statements are to be added that corroborate the hypothesis, with which it is inductively consolidated. Thus, verificationism clashes with falsificationism in the function assigned to observation statements. From falsificationism's viewpoint, no scientific theory can ever be deduced from observation statements or can be described as a truth function of observation statements.

In distinction from both of the foregoing approaches, confirmation holism (also called *epistemological holism*) is skeptical of the very possibility of testing a theory. Confirmation holism states that it is impossible to test a scientific hypothesis in isolation, since an empirical experiment requires assuming as true one or more auxiliary hypotheses (or background assumptions). The hypothesis in question is itself incapable of making predictions. In fact, the consequences of the hypothesis typically are based on the auxiliary hypotheses from which predictions derive. This argument prevents a theory from becoming conclusively falsifiable through the empirical world if auxiliary hypotheses are not tested, which is an improbable task since auxiliary hypotheses sometimes include one or more scientific theories.

Because this thesis was formulated by Pierre Duhem (1861–1916) and Willard Van Orman Quine (1908–2000), it is also known as *Duhem–Quine thesis*. Its implications for the possibility of confronting a theory with evidence are critical. Because a single hypothesis or theory can only be tested by its relation to other hypotheses or theories, this leads us to the problem against which confirmation holism warns us: One might believe that the results of a test have refuted a hypothesis or scientific theory when actually these results have not done so at all. Instead, it can be argued that those results conflict with the predicted effect because any other hypothesis or theory is false. This argument prevents a theory from becoming conclusively falsified.

From a distinctive epistemological outlook that contrasts with those of falsificationism, verificationism, and confirmation holism, pragmatism offers what this tradition calls the *pragmatist maxim*. This can be defined as a rule for clarifying the contents of hypotheses by tracing their practical consequences. Instead of *true* theories, pragmatism prefers to speak of theories that perform

as if they were true. True propositions, then, are those that enable us to function well. True statements are ones that function successfully as instruments. From pragmatism's viewpoint, propositions are evaluated by the merits of the practices that are involved in our use of them.

Consequently, confronting a hypothesis or theory with evidence involves identifying its practical consequences. Pragmatism argues that the *clarification* of a scientific hypothesis implies we have the information we need for testing it empirically. Testing theories involves carrying out experiments, understood as the performing of rational actions. If the hypothesis is not true, then the experiment will fail to have some predetermined practical effect. If the hypothesis is wrong, we will observe *practical differences* between our theoretical expectations and what the experiment shows us.

Falsificationism proposes R as criterion of demarcation; R distinguishes science from pseudoscience. Contrary to the requirement of searching for elements that refute a theory, the most characteristic element of pseudoscience is the incessant stream of confirmations—that is to say, of observations that *verify* the theories in question. Pseudoscience always involves confirmatory information. Whatever happens always confirms it. Thus, its truth appears to be manifested. Classic examples of supposed pseudosciences are three theories: Karl Marx's theory of history, Sigmund Freud's psychoanalytic theory, and Alfred Adler's individual psychology. According to the characteristics of pseudoscience, every conceivable case (or observation) could be interpreted in light of Adler's theory, or of the Marxist theory, or equally of Freud's theory. It would be difficult, if not impossible, to think of any social or human behavior that could not be interpreted in the terms of any of these theories.

This idea may be illustrated by a classic example of *self-verification* for both the Adlerian theory and Freudian theory through their *explanations* of two kinds of human behavior, which drew attention to the philosopher Karl Popper (1902–1994): that of a man who pushes a child into the water with the intention of drowning the child; and that of a man who sacrifices his life in an attempt to save the child. Each of these two cases can be explained with equal ease in Freudian and in Adlerian terms. According to Freud, the first man suffered from

psychological repression, while the second man achieved what Freud called *sublimation*. According to Adler, the first man suffered from feelings of inferiority (perhaps the need to prove to himself, in that he dared to commit a crime), and so did the second man (whose need was to prove to himself, in that he dared to rescue the child). It is precisely this feature—that these theories always are confirmed—that constitutes its logical weakness. Different observation statements lead to confirmation of the same theory.

As an example of science (and a counterexample of the pseudosciences), Popper mentions the endeavored attempt to refute just one prediction of the theoretical physicist Albert Einstein's gravitational theory, made by Arthur Eddington's 1919 expedition. Einstein's hypothesis stated that light was attracted by heavy bodies (such as the sun), precisely as material bodies were attracted. The prediction was that light from a distant, fixed star, which had an apparent position that was close to the sun, would reach the earth from such a direction that the star would seem to be slightly shifted away from the sun. This is something that cannot normally be observed, since such stars are rendered as invisible in the daytime by the sun's overwhelming brightness; however, during an eclipse, it is possible to take photographs of them. If the same constellation is photographed at night, one can measure the distances in the two photographs and check for the predicted effect. The findings of Eddington confirmed Einstein's gravitational theory or, rather, one of the hypotheses of that theory.

The central issue is that Einstein's theory was incompatible with certain possible results of observation. If observations showed that the predicted effect was definitely absent, then the theory would simply be refuted. Every scientific theory is a prohibition: It forbids certain occurrences from happening. The more a theory forbids, the better it is. This is quite different from the pseudoscience that was previously described: When the theories in question are compatible with the most divergent human behavior, it is practically impossible to describe any human behavior that might not be claimed to be a verification of these theories.

A second issue is that only the theory of Einstein—in opposition to the Freudian theory, Marxist theory, and Adlerian theory—was heavily exposed to a high

risk of being refuted. An important disadvantage for a scientific theory is that it is easy for a nonscientific theory to be *confirmed* or *verified*. This occurs for almost every theory, if their adherents look for confirmations. At the same time, there are degrees of testability. Some theories are more testable, and some theories are more exposed to refutation than others; they involve, as it were, greater risks.

Some definitely testable theories, even when they are found to be false, are still upheld by their followers. The followers rescue the theory from R through some *procedures*. Examples of these procedures are introducing ad hoc some auxiliary assumption or re-interpreting the theory ad hoc in such a way that it avoids R. These decisions, however, have a fundamental cost for the theory: destroying, or at least decreasing, its scientific status.

The Marxist theory of history, despite the serious efforts of some of its founders and followers, ultimately adopted those practices. In some of its earlier formulations (e.g., in Marx's analysis of the character of the *coming social revolution*), the predictions were testable and were in fact falsified. Instead of accepting the refutations, however, Marxists reinterpreted both the theory and the evidence in order to make them agree.

The Freudian and Adlerian theories were in a different class. They were simply nontestable and irrefutable. There was no conceivable human behavior that could contradict them. Typically, supporters of both theories reinterpreted their hypotheses ad hoc in such a way that they escaped refutation. This does not mean that both theories did not present valid viewpoints; however, it does mean that the *clinical observations*, which some analysts naively believe confirm their theory, cannot supply such confirmations.

These considerations lead to the problem of establishing a criterion of demarcation among the statements or systems of statements for the empirical sciences, as well as all other statements, such as religious, metaphysical, or pseudoscientific statements. The criterion of falsifiability was the Popperian solution to the problem of demarcation: statements or systems of statements, in order to be defined as scientific, must be capable of conflicting with possible or conceivable observations. In a different way, the criterion of verifiability claims that only those propositions that are deducible from true observation statements belong to the science.

Consequently, from the verificationism's viewpoint, the meaning of confronting a theory with evidence is established through true observation statements that can verify a hypothesis or theory.

Against falsificationism's position in favor of refutation (i.e., the criterion for testing a theory must be to prove that this theory is wrong or false), confirmation holism poses a criticism to Popper and his criterion of falsifiability: As the Duhem–Quine thesis argues, for any alleged falsification it's always possible to maintain a hypothesis by revising its auxiliary hypotheses.

Pragmatism also has implications in terms of the demarcation problem. A pragmatic approach to the line, or criterion, of demarcation would argue that while there are some core principles that one can use in distinguishing between science and nonscience, judgments and decisions about the scientific status of a hypothesis or theory depend, in part, on practical goals and concerns. In a manner distinct from other views on the demarcation problem, pragmatism claims in favor of considering both theoretical statements and practical statements as lines to distinguish science from not-science.

Induction, Conjectures, and Refutation

Following the Scottish philosopher David Hume's (1711–1776) criticism of the logical basis of induction, Popper was led to his theory of conjectures and refutations (or theory of trial and error). Popper's thesis—following Hume's ideas on induction—asserted that our attempts to force interpretations upon the world were logically made prior to the observation of similarities. This, in turn, had an important consequence for the philosophy of science; scientific theories were not the digest of observations, but they were inventions—that is, conjectures boldly presented for trial and to be eliminated, if they clashed with observations.

In this sense, Hume gives a causal explanation about the fact that we believe in laws or in statements that assert regularities or constantly conjoined events, by asserting that this fact is due to a *custom* or *habit*—that is, the custom or habit of believing in laws or regularities. It can therefore be said that, like other habits, our habit of believing in laws is the product of frequent repetition of the repeated observations that

things of a certain kind are constantly conjoined with things of another kind. The central idea of Hume's theory is repetition, based on similarity (or *resemblance*).

In the wake of Hume's theory, Popper noted that only repetition-for-us, based upon similarity-for-us, can have any effect upon us. We must respond to situations *as if* they were equivalent, viewing them as similar and interpreting them as repetitions. Thus, the cases were repetitions only from a certain point of view. This meant that, for logical reasons, there must *always* be a point of view, such as a system of expectations, anticipations, assumptions, or interests *before* there can be any repetition; point of view, consequently, cannot be merely the result of repetition.

However, if this is so, there can be no valid logical arguments that allow us to establish that the instances, of which “we have had no experience,” resemble instances of which “we have had experience.” Consequently, even after observation of the frequent or constant conjunction of objects, we have no reason to draw any inference regarding any object beyond instances of which we have had experience. In other words, an attempt to justify the practice of induction by an appeal to experience must lead to an infinite regress. As a result, we can say that theories can never be inferred from observation statements or rationally justified by them.

Having refuted the logical idea of induction, Hume was faced with the following question: how do we actually obtain our knowledge if induction is a procedure that is logically invalid and rationally unjustifiable? There is only one possible answer: we obtain our knowledge by a noninductive procedure. Conjectures, counterexamples, and refutations were solutions that falsificationism proposed for the tricky basis of induction.

A *conjecture* is a proposition that is unproven. A conjecture stands in contrast to a *hypothesis*, which is a testable statement based on accepted grounds. In mathematics, a conjecture is an unproven proposition that appears to be correct. Sometimes, a conjecture is called a *hypothesis* when it is used frequently and repeatedly as an assumption to prove other results. For example, the Riemann hypothesis is a conjecture from number theory that (among other things) makes predictions about the distribution of prime

numbers. Few number theorists doubt that the Riemann hypothesis is true.

In mathematics, any number of cases that support a conjecture, no matter how large, is insufficient for establishing the conjecture's veracity, since a single counterexample would immediately invalidate the conjecture. Formal mathematics is based on *provable truth*.

In terms of logic, especially in its application to mathematics and philosophy, a *counterexample* is an exception to a proposed general rule. For instance, consider the proposition that "all apples are small." Because this statement makes the claim that a certain property (smallness) holds for all apples, even a single example of a big apple will prove it to be false. Thus, any big apple is a counterexample to "all apples are small." More precisely, a counterexample is a specific instance of the falsity of a universal quantification (a "for all" statement).

Conjectures disproven through counterexamples are sometimes referred to as *false conjectures*. A classical example of a false conjecture is the Pólya conjecture. Regarding number theory, the Pólya conjecture stated that *most* (that is, 50% or more) of the natural numbers, less than any given number, have an odd number of prime factors. The conjecture was posited by the Hungarian mathematician George Pólya in 1919 and was proven as false in 1958 by C. Brian Haselgrove. The size of the smallest counterexample is often used to show how a conjecture can be true for many numbers and still be false.

Refutation indicates that a scientific theory in a *conjecture process* must be proven as valid. Clearly, refutation and falsifiability are related concepts. Often, they are used as interchangeable notions. However, there seems to be a subtle difference between the two. One does not need to have an observationally testable theory to be able to effectively refute a scientific theory. Refutation and criticism are the basis for all gains in knowledge. According to Popper, we *only gain knowledge from our mistakes*. It is, therefore, *not* observation that initiates the growth of knowledge. Thus, contrary to the inductivist scheme, scientific discovery does not need to begin with observational evidence. In other words, scientific enterprise always advances in relation to a set of ideas that seems to be inadequate and worth trying to improve.

Refutation, Falsification, and Verisimilitude

Falsification and verisimilitude concepts play roles in the notion of truth, an uneasy idea that Popper avoided in his earliest writings. Instead, it was less risky to accept that every theory is an open-ended hypothesis. Then, *ipso facto* it must be at least potentially false. For this reason, it was more acceptable to assert that a theory that is falsified is false and is known to be such, as well as a theory that replaces a falsified theory (because it has a higher empirical content than the latter and presents what has falsified it), is a *better theory* than its predecessor. However, Popper came to accept Alfred Tarski's (1901–1983) reformulation of the correspondence theory of truth, and he integrated the concepts of truth and content to frame the metalogical concept of *truthlikeness* or *verisimilitude*. Thus, Popper argued that a *good* scientific theory has a higher level of verisimilitude than its rivals. He also explicated this concept by referring to the logical consequences of theories.

Scientists should therefore be interested in theories with highly informative content because such theories possess a high predictive power and are, consequently, highly testable. However, if this is true, then, as paradoxical as it may sound, the more *improbable* a theory is, the better it is scientifically. This is because the probability and informative content of a theory vary inversely. The higher the informative content of a theory, the lower its probability will be; the more information a statement contains, the greater the number of ways it may turn out to be false. Informative content, which has an inverse proportion to probability, is in direct proportion to testability.

The Popperian view proposes two methods of comparing theories, in terms of verisimilitude, the qualitative and quantitative accounts. In the qualitative approach, it is asserted that assuming that the truth-content and the falsity-content of two theories t_1 and t_2 are comparable, we can say that t_2 is more closely similar to the truth or corresponds better to the facts than t_1 , if and only if either: (1) the truth-content but not the falsity-content of t_2 exceeds that of t_1 ; or (2) the falsity-content of t_1 but not its truth-content exceeds that of t_2 .

Here, verisimilitude is defined in terms of subclass relationships: t_2 has a higher level of

verisimilitude than t_1 , if and only if their truth (and falsity-content) are comparable through subclass relationships and if either of the following are true: (1) t_2 's truth-content includes t_1 's and t_2 's falsity-content, if it exists, is included in, or is the same as t_1 's; or (2) t_2 's truth-content includes or is the same as t_1 's and t_2 's falsity-content, if it exists or is included in t_1 's.

In the quantitative approach, verisimilitude is defined by assigning quantities to contents, and the index of the content of a given theory is its logical improbability (given again that content and probability vary inversely). Formally, then, Popper defines the quantitative verisimilitude, which a statement, "a," possesses by means of the formula:

$$Vs(a) = Ct_T(a) - Ct_F(a),$$

where $Vs(a)$ represents the verisimilitude of a , $Ct_T(a)$ is a measure of the truth-content of a , and $Ct_F(a)$ is a measure of its falsity-content.

The utilization of either method of computing verisimilitude shows that even if theory t_2 , with a higher content than a rival theory t_1 , is subsequently falsified, it can still legitimately be regarded as a better theory than t_1 . Here, *better* is understood as meaning that t_2 is *closer* to the *truth* than t_1 . Thus, scientific progress involves, in this view, the abandonment of partially true but falsified theories.

Lakatos's view on scientific progress contrasts with Popper's argument in favor of verisimilitude. Imre Lakatos claims that merely defining the aim of science as verisimilitude is not enough. One needs to be able to identify this progress when it occurs. It is not enough to say that novel corroboration may be a "good reason to conjecture" that we are getting closer to the truth. We need a firm connection between the two.

On the other hand, the role of R in scientific progress is completely different for these two philosophers of science. In Lakatos's terms, scientific progress is defined as predicting some *novel fact*—that is, a *hitherto unexpected fact*. According to Popper, novel facts are a bit of an add-on, since much of the empirical burden is carried by falsification and refutation. For Lakatos, who wanted to decrease (to zero) the role of negative evidence and refutation, the entire burden of scientific progress is carried on the shoulders of novel facts.

Lakatos asserts that the importance of novelty is a clear counterexample of scientific progress; if we really know P1: "Swan A is white," then Pw:

"All Swans are white" represents no progress because this may only lead to the discovery of similar facts, such as "P2: Swan B is white." Hence, so-called *empirical generalizations* constitute no progress. A *new* fact must be improbable or even impossible in the light of previous knowledge.

From the Popperian view, scientific theories might change in the future, as we find an observation that refutes existing theory. However, until a new fact comes into existence, we are justified in believing in the knowledge gained through the conjecture and refutation process. Induction is neither needed nor desired in this formula for growth of knowledge.

Finally, conjecture and refutation are not merely the ways in which science progresses. For example, Popper sees direct ties between evolution through natural selection and the growth of scientific knowledge. In fact, growth of scientific knowledge may be thought of as a special case of biological evolution. From a biological or evolutionary point of view, science or progress in science may be regarded as a means used by the human species to adapt to an environment, to invade new environmental niches, and even to invent new environmental niches. Therefore, all of the knowledge, even the *knowledge* embodied in DNA, can be conceived as stemming from the conjecture-and-refutation process.

Rodolfo Sarsfield

See also Consistency; Falsifiability; Hypothetico-Deductivism; Inferentialism; Soundness

Further Readings

- Kelly, L. (2013, July 14). The Duhem–Quine thesis reconsidered—Part one. *Critical Rationalism Blog*. Retrieved April 2017 from <https://www.criticalrationalism.net/2013/07/14/the-duhem-quine-thesis-reconsidered-part-one/>
- Hands, D. W. (1993). *Testing, rationality, and progress. Essays on the Popperian tradition in economic methodology*. Rowman and Littlefield.
- Hookway, C. (2016). Pragmatism. In E. N. Zalta (Ed.), *The Stanford encyclopedia of philosophy* (Summer 2016 Edition). Retrieved April 2017 from <https://plato.stanford.edu/archives/sum2016/entries/pragmatism/>
- Hume, D. (2000). *A treatise of human nature: Being an attempt to introduce the experimental method of reasoning into moral subjects* (3 Vols). Oxford University Press. (Original work published 1739–1740)

- Lakatos, I. (2000). *Falsification and the methodology of scientific research programs*. In T. Schick (Ed.), *Readings in the philosophy of science: From positivism to postmodernism*. Mayfield.
- Popper, K. R. (1963). *Conjectures and refutations: The growth of scientific knowledge*. Routledge.
- Popper, K. R. (1976). A note on verisimilitude. *British Journal for the Philosophy of Science*, 27, 147–159.
- Popper, K. R. (1994). *The myth of the framework: In defense of science and rationality*. Routledge.
- Stein, S. K. (2010). *Mathematics: The man-made universe*. Dover.
- Thornton, S. (1914). Karl Popper. In E. N. Zalta (Ed.), *The Stanford encyclopedia of philosophy* (Summer 2014 Edition). Retrieved July 2014 from <https://plato.stanford.edu/archives/sum2014/entries/popper/>
- Wittgenstein, L. (2001). *Tractatus logico-philosophicus*. Routledge. (Original work published 1922)

RELATIVE CONSISTENCY

Relative consistency is a theoretical construct the utility of which arises mainly in the fields of mathematics and logic, although its use has been extended (through the use of metaphor and analogy) to other, more loosely defined contexts. At its most basic level, relative consistency is a relation that may hold between different components of a theory. It will sometimes be the case that a theory cannot be objectively proven; in most cases, this will also mean that the individual components of the theory also cannot be proven, either because no objective standard or method of verification exists, or because it does exist but is unavailable. In such cases, it may still be possible to show that the individual parts of the theory are consistent with one another, or *relative to* one another. If this can be done, then, as shown in the following sections, one can prove the relative consistency of the theory even if the theory itself cannot be objectively proven.

Consistency

A theory is considered to be consistent if it is free from any contradictions. There are two main methods of determining whether or not a theory contains one or more contradictions. The first method is syntactic, and the second method is

semantic. The *syntactic method* asks us to determine that there is no formula that can be constructed using the theory's axioms such that both the formula and its negation can be proven. If there is, then the theory is not consistent, but if there is not, then the theory is consistent under syntactic analysis.

Semantic analysis of a theory's consistency asks us to determine whether there is a model of the theory within which all of the theory's formulas are demonstrably true. If it is possible to show such a model, then the theory is semantically consistent. If no such model can be shown, then the theory is not semantically consistent.

Within the realm of mathematics, the method used to determine whether a theory is consistent is usually referred to as a *consistency proof*, a special type of the broader category of mathematical proofs.

Completeness

The completeness of a logical or mathematical theory is intricately bound up with its consistency. A theory is said to be complete if the theory itself is consistent but none of the theory's extensions is consistent. One way that this can be determined when one is working in the realm of classical logic is by examining each sentence that can be composed in the language of the theory. If each of these sentences, or its negation, exists within the theory already, then the theory meets the criteria for completeness.

Gödel and Incompleteness

Kurt Gödel, an Austrian mathematician who lived from 1906 until 1978, was active in the fields of logic and philosophy and is best known for his work on the concept of incompleteness. Essentially, Gödel was able to demonstrate that any theory of arithmetic that is complex enough will be, in a sense, incomplete at the most basic level because it will not be possible for the theory to be both consistent and complete—only one or the other will be attainable. The reason for this is difficult to explain without lengthy mathematical proofs, but it has to do with the concept of *recursion*, also known as *self-reference*. A sufficiently complex theory can never be both consistent and

complete because in order to be complete, it would need to contain a full copy of itself. This seems simple enough until we recall that inside that copy of itself there will need to be another copy of the theory itself, nested within the parent copy—otherwise the copy would be incomplete. This paradoxical dilemma will continue to replicate itself downward, layer upon layer, without end, never attaining completeness.

Some who have sought to carry on Gödel's work have attributed to it a semimystical significance, suggesting that it is symbolic of the ultimately ineffable nature of the universe, and affirming the suspicions of mystics that there are some things in the universe that are simply unknowable. That point is certainly open for debate, but other applications of Gödel's work are not in controversy. The most salient takeaway from Gödel's work on incompleteness is that when an arithmetical system—or a set of theoretical models that has been translated into mathematical symbols—is complex enough, it will be impossible to prove the consistency of that theory on its own terms. This is because the system of the theory will always be incomplete (due to the recursion problem noted earlier), and therefore, the consistency of the complete system is not possible to verify. This creates special problems for mathematicians and logicians seeking to establish the consistency of a theoretical model. We would like to establish models that are both complete and consistent, yet Gödel predicts that this is impossible using conventional definitions of what it means to be complete and consistent. The solution to this dilemma, when it eventually appeared, was both simple and elegant: create and apply a new definition of consistency that, while different in significant respects from the traditional, objective concept of consistency, is still persuasive evidence of a theoretical model's utility. This type of consistency is known as *relative* because its focus is not on how objectively consistent a theory is but on whether it is consistent with another theory already known to be consistent.

The Value of Relative Consistency

At this point, it should be clear why relative consistency is of interest in the analysis of mathematical and logical theories: Because it is not possible

to demonstrate consistency of sufficiently complex theories, we instead evaluate them for relative consistency.

It is useful to recall from the study of logic that a formal system is inconsistent if it is possible for us to use the system to prove both an axiom and the negation of that axiom. As Gödel made clear, it can be extremely challenging to prove that a system is consistent, given the potentially very large number of statements that can be formulated using the language of that system, and given the essentially incomplete nature of very complex systems. What is often easier to do is to prove that a system is consistent in relation to another system, the consistency of which has already been verified.

For example, if one were to develop a system of axioms A, and there also exists a system B, the theorems of which are the same as the axioms of A, then by comparing the two, we could determine that A and B are relatively consistent, that is, they are consistent relative to one another rather than consistent with an objective standard. A historical example of this dynamic can be seen in the case of the Russian mathematician Nikolai Ivanovich Lobachevski (1792–1856). Lobachevski was the first to publish his research into hyperbolic geometry (also referred to as *non-Euclidean geometry*, or simply *Lobachevskian geometry*). Lobachevski's geometry is congruent with Euclidean geometry except for one point. Where Euclid's fifth postulate states that, for a given line and a point not on the line, there is only one line that passes through the point not on the line but does not intersect with the line, Lobachevski developed a geometry in which more than one line meets these requirements. Lobachevski's work was validated in part through the application of principles of relative consistency.

This process is known as a *relative consistency proof*. Relative consistency proofs can become extremely complicated, since they require one to create a theory interpretation into another theory, so in recent years, there has been a move toward accomplishing this task with the use of computer programs whenever possible. This requires the user to enter the axioms of the theories being compared into mathematical statements capable of being manipulated by the computer software. Once this has been done, the computer can

evaluate the axioms for consistency relative to one another by breaking the theories down into simpler subsystems, analyzing these, and then “putting the pieces back together.”

Relative consistency is widely used as an alternative to proving completeness or consistency. Gödel himself used relative consistency to show the independence of the Axiom of Choice, a part of set theory, and the same method was also used to prove the Continuum Hypothesis, which concerns the sizes possible for infinite sets of numbers. Gödel’s career was unusual inasmuch as he not only became aware of the nature of a huge obstacle to progress in logical and mathematical theory—incompleteness—but he also developed an ingenious workaround for this obstacle using relative consistency to force consistency proofs to “pull themselves up by their bootstraps,” in other words demonstrating consistency toward each other even when objective consistency could not be established.

Applications of Relative Consistency

As mentioned earlier, relative consistency had its conceptual origins in mathematics and logic, but it has been adapted to other fields and incorporated into the vernacular of everyday speech. From its original, technical meaning, it has evolved to the point where it is often used to describe the extent to which a set of results make sense when considered together. For example, if we were to collect test scores from students at a high school and analyze them, without even being aware of it, we would be checking for discrepancies—places where particular parts of the data do not make sense when compared with other parts. We might find that over 80% of students scored no higher than a C in algebra but that over 90% of students received a grade of B or an A in trigonometry. If either the algebra grades or the trigonometry grades were viewed in isolation, there would be nothing remarkable about them to draw our attention. When they are considered in relation to one another, however, they appear to be inconsistent because trigonometry is traditionally seen as a much more difficult topic than algebra. Assuming that is the case here, why would most people do worse in the easier class than they did in the more difficult one?

This raises a related use for relative consistency concepts outside of logic and math: test construction and validity. When authors set out to create an assessment instrument to be administered to a population of subjects in order to measure one or more qualities the subjects may possess, their challenge is to create questions that actually test the attribute the designers intend to assess rather than some unrelated characteristic. One of the most important methods used to ensure that this happens is to review the results of the instrument for relative consistency. An interesting example of a question that tested something other than had been intended (the true target was students’ knowledge of vocabulary) involved a multiple-choice question asking students which would be more appropriate to go fishing on: a yacht, a schooner, a canoe, or a catamaran. Undoubtedly, the person who drafted the question thought that it was a perfectly reasonable way to assess students’ nautical terminology. However, when the test administrators noticed that students in certain parts of the country tended to score poorly on this question, issues of relative consistency were raised, and further study was devoted to understanding the reason for the discrepancy. It turned out that students from lower income neighborhoods and students who lived in landlocked states had very little knowledge about different types of watercraft and thus were unable to discern the correct answer to the question. Thus, the question was really testing a type of knowledge that only a certain type of student would have rather than knowledge all students could reasonably be expected to have. As in the previous example, if the question had been reviewed in isolation, there would have been no way of understanding why it was being answered incorrectly. But, by viewing it in relation to students’ performance on other questions, test administrators were able to identify the issue responsible for the drop in scores on that question.

Scott Zimmer

See also Induction; Logic, Formal and Informal; Mathematics, 20th century

Further Readings

Dawson, J. W. (1997). *Logical dilemmas: The life and work of Kurt Gödel*. A. K. Peters.

- Machamer, P. K., & Turnbull, R. G. (1976). *Motion and time, space and matter: Interrelations in the history of philosophy and science*. Ohio State University Press.
- Quandt, J. (1987). Relative consistency of a set theory with hyperclasses. *Mathematical Logic Quarterly*, 33(2), 101–106.
- Ranney, M. (1994). Relative consistency and subjects' "theories" in domains such as naive physics: Common research difficulties illustrated by Cooke and Breedin. *Memory & Cognition*, 22(4), 494–502.
- Shapiro, S. (2014). *Varieties of logic*. Oxford University Press.
- Son, L. K., & Vandierendonck, A. (2007). *Bridging cognitive science and education: Learning, memory and metacognition*. Psychology Press.
- Wisse, P. (2001). *Metapattern: Context and time in information models*. Addison-Wesley.

S

SCIENTIFIC REALISM

Scientific realism (SR) constitutes a variety of realism that emerged primarily with the work of theorists Karl R. Popper and Mario Bunge. SR asserts that a real external world exists and that we can know it through the scientific method. SR holds a strong commitment to the idea that scientific theories have a privileged epistemic status compared with other forms of knowledge such as ordinary knowledge, intuition, or hermeneutics. SR also maintains that scientific theories yield knowledge of unobservable aspects of the world. The basic principles of SR are as follows: (1) The external world exists independently of the cognitive subject (ontological realism); (2) it is possible to know the world (epistemological realism); (3) some propositions refer to facts and some factual propositions are approximately true (semantic realism); (4) all factual knowledge is incomplete and fallible (fallibilism); (5) in principle, any factual proposition is perfectible (meliorism); (6) the best approach to explaining the world is the scientific method (methodological realism). These principles are discussed in the sections that follow.

SR has similarities and differences with two other varieties of realism: *naïve realism* and *critical realism*. Naïve realism accepts principles (1) and (2). Critical realism accepts principles (1), (2), (3), (4), and (5). SR holds the principles (1), (2), (3), (4), (5), and (6). A distinctive attribute of SR over other forms of realism is its methodological realism. Methodological realism holds that the

best approach to study entities, forces, and processes, whether observable or unobservable, is to use the scientific method. Using the scientific method, we can reasonably construe scientific theories as providing knowledge. This is true for both natural and social facts.

SR is an amalgamation of two rival classical philosophical traditions: rationalism and empiricism. SR promotes the construction of theories competent to account for empirical data, as well as the exploration of data in light of theory. Hence, it discourages unrestrained speculations as well as mindless data gathering. SR emphasizes the constructive nature of ideas in particular concepts, problems, rules, and experimental designs. It stresses the justification—conceptual, empirical, or both—of all claims to knowledge. However, it is fallibilist as well as meliorist. Namely, it recognizes no final truths except in logic and mathematics. Lastly, SR is systemic: It places every question and every answer into some system, such as description, classification, theory, or a well-designed empirical research project.

Realism and Antirealism

Antirealism rejects one or the other of these two realist theses: (1) The external world is material and exists by itself and (2) science can capture that reality. Roughly speaking, antirealism holds that scientific explanation is inherently limited because science seeks only relations and does not inquire into the nature of objects, or because its epistemological scope is fundamentally limited. Moreover,

some antirealists contend that claims about unobservable things have no literal meaning at all. There are, however, several kinds of antirealism.

Transcendentalism adds that every entity has an immutable essence beneath its properties and relations and that those essences, which escape reason as well as experience, are apprehensible with nonscientific means: with intuition of some sort or other. *Phenomenalism* refuses to tolerate essences just because they are nonphenomenal, but it agrees with transcendentalism that science is limited to establishing relations, whether among phenomena or among fairly directly observable traits or facts, but that it never gets to the *things themselves*, which are mere *metaphysical constructs*. Realism holds that experience is not definitive but should be accounted for in terms of a much wider but indirectly knowable world: the set of all existents.

Phenomenalism, by contrast, asserts that the endeavor of science consists of collecting, describing, and systematizing phenomena without inventing diaphenomenal or transobservational objects—such as quarks, mental representations, and social cognition. For realists, experience is a type of event: Every single experience is a fact happening in the cognitive subject, which is in turn regarded as a concrete system (e.g., a brain)—but one having expectations and a set of knowledge that both alter and enrich their experiences. Therefore, realism encourages the development of theories going beyond the systematization of experiential items and requiring consequently ingenious test procedures.

SR rejects all types of irrationalism, from neo-Platonism to existentialism and *postmodernist* relativism–constructivism, passing through intuitionism and pragmatism. Against these philosophies of science, SR asserts that no knowledge is possible without the combination of reason and experience.

SR and Other Schools of the Philosophy of Science

The philosophy of science in SR is a sort of synthesis of realism, rationalism, and empiricism. SR asserts, as does realism, that there is an external world, that scientific research aims at depicting it, and that this is a never-ending enterprise because every picture of the world misses some of its traits

and adds others that are untrue. Simultaneously, SR claims, as does rationalism, that such pictures are symbolic rather than iconic, that logic and mathematics are a priori, and that signs are meaningless unless they stand for ideas. Finally, SR affirms, as does empiricism, that experience is a crucial feature for factual theories and procedures, that there can be no certainty in them, and that philosophy should adopt the method and standards of science. Therefore, SR marries its position to the successful reference of theoretical terms, including those for unobservable entities, processes, properties, and relations.

SR is an alternative to other major traditions in the philosophy of science. SR clashes with radical rationalism as well as irrationalism, with intuitionism and pragmatism, with classical positivism and logical positivism, and with constructivism–relativism as well as critical rationalism.

The radical rationalist believes we come to knowledge “a priori”—through the use of logic—and that this knowledge is, therefore, independent of sensory experience. In a different manner, SR emphasizes that experience with the external world is another source of human knowledge.

Irrationalism argues in favor of two theses in opposition to SR. First, irrationalism claims that intuition, instinct, feeling, or faith, rather than reason are the sources of human knowledge (epistemological irrationalism). Second, irrationalism holds that the universe is governed by irrational forces (ontological irrationalism). In contrast, SR asserts that reason (and not intuition, instinct, or faith) is the mental basis for scientific knowledge. Irrationalism, according to SR, is a philosophy that seeks reasons to justify irrationality, which is a contradiction on its own terms. On the other hand, and contrary to ontological irrationalism, SR asserts that the assumption that nature and culture are irrational has been largely refuted by the advancement of science. All of these arguments can also be said about the differences between SR and intuitionism.

Pragmatism claims that thinking about problems should be based on dealing with a specific situation instead of on ideas and theories. In other words, pragmatism asserts that the meaning of concepts is to be sought in their practical bearings, that the function of thought is to guide action, and that truth is preeminently to be tested by the

practical consequences of beliefs. There is an epistemological difference between SR and pragmatism: according to pragmatists, prediction is all-important, whereas explanation is parasitic. Prediction is more valid than explanation in epistemic terms. A second difference between SR and pragmatism is about the aim of science: For pragmatists, the maximum value of a theory is its efficiency in relation to practical problems, whereas for a scientific realist the maximum value of a theory is to produce true descriptions of things in the world. Since theories can be applied to cognitive as well as to practical aims, therefore, SR and pragmatism are not two opposing enterprises. However, SR highlights that cognitive application of theories (e.g., explanation) precedes their practical applications (e.g., technologies).

Classical positivism, defined as positivism prior to the foundation of the group of philosophers known as the *Vienna Circle* in the early 1920s, holds that impressions of the senses provide direct knowledge of the world. A response from SR to classical positivism is that, for instance, the light provided by a source such as the sun does not indicate in a direct manner the composition and structure of the source. The light of the sun tells us nothing about the fact that the sun is composed mostly of hydrogen. Likewise, the light of the sun informs us neither that the hydrogen atom consists of a single electron, nor that the electron can be in any of infinite states, nor that the change from one state to another generates a photon. Therefore, the perception of light through the eyes and the brain does not tell us, by itself, anything about its composition. Such composition and structure are unobservable, and they need to be explained by a theory (provided by reason and not only by perception).

A second difference between classical positivism and SR is about the role of data in scientific inquiry. Classic positivism affirmed that data guide the whole sphere of scientific research. In a different manner, SR claims that data are neither used nor organized in any way, but also that a set of hypotheses is always used in a more or less explicit manner. Usually, the data are sought and produced in accordance with some hypotheses. These, in turn, are controlled by relevant empirical evidence. Neither the data nor the hypotheses have the final word in scientific research. What it

is true is that no perfect—completely exhaustive and accurate—explanation is ever achieved. Consequently, there are no final explanations.

The logical positivists, separating from their predecessors, claim to have overcome the dilemmas of realism/antirealism and materialism/idealism. In particular, they argue that the autonomous existence of an external world is a pseudo-problem. On the other hand, and like SR, most logical positivists reject religion and metaphysics. SR and logical positivists agree in considering that the physical world is either possibilities of perception or logical constructions derived from precepts. Logical positivism also adheres, however, to some of the theses of phenomenism. In particular, some of its members argued that the Copernican system of cosmology differed from the Ptolemaic system only in that it was another way of speaking. The major dispute between these two opposing theories about the organization of the planetary system is seen just as a problem of language on the part of the logical positivist. Whatever the reality was, it passed to the sidelines. This view places logical positivism far away from SR.

Critical rationalism, represented by the work of the Anglo-Austrian theorist Karl R. Popper (1902–1994), tries to assemble a theory of the proper assessment of sentences, that is, of the opportunity of proving the truth or falsity of some sentences. Popper begins with the fact that a theory is false if it contradicts a singular sentence describing some observation reports. Popper then affirms that such singular sentences are veridical, that is, truthful as opposed to illusory, so they may be used to create ultimate proofs of the falsity of some universal sentences. For instance, the singular sentence, “That swan is black,” if it is a true report of some observation, can be used to generate a final proof of the falsity of the universal sentence, “All swans are white.” But, he argues, proof of universal sentences or the demonstration that they are probable requires inductive inferences.

Inductive inferences have observations as premises and theories as conclusions. They are notoriously invalid but often are deemed unavoidable. As a consequence, no such putative proof can be valid. SR’s strong version, exemplified by the Argentine philosopher and physicist Mario Bunge’s oeuvre, apparently finds more affinity with those thinkers who emphasize the use of

formal methods and who futilely seek justification, than with those who deny the possibility of justification and deem the use of formal methods more limited than he does. As a corollary of this attitude, Bunge proposes to respond to difficulties first with small changes that preserve systems (or frameworks) and move to larger ones when these prove inadequate. This is the exact opposite of what Popper said, as he advocated that one should always prefer that theory which has the highest degree of falsifiability. These thinkers, however, disagree not about the aim of the philosophy of science, which is to improve critical standards so that the best possible theories are created and honed, but rather about the best means for doing that.

The constructivist-relativists deny the difference between thing and idea, between fact and fiction, and between law and convention. They deny the existence of objective truths and, therefore, they oppose the existence of transhistorical and transcultural truths. Conversely, SR holds that ideas and facts are different entities and that they exist independently of the observer. The facts are not confined to constructions of researchers. The most notable example is provided by mathematics, chemistry, and genetics: Their truths are independent of the culture and its conventions.

SR, Naive Realism, and Critical Realism

When referring to SR, scholars argue, more or less consistently, the following postulates: (1) Science aims to produce true descriptions of things in the world (or approximately true descriptions); (2) the best existing scientific theories are at least approximately true; (3) the central sentences of the best current theories are genuinely referential; (4) the approximate truth of a scientific theory is sufficient explanation of its predictive success; (5) the (approximate) truth of a scientific theory is the only possible explanation of its predictive success; (6) a scientific theory may be approximately true even if inferentially unsuccessful; (7) the history of at least the mature sciences shows progressive approximation to a true account of the material world; (8) the theoretical terms of scientific theories are to be read literally, and so read are definitively true or false; (9) scientific theories make genuine, existential claims; and (10) the predictive

success of a theory is evidence for the referential success of its central terms.

It is important to distinguish between a strong (or comprehensive) version and a weak (or restrained) version of SR. For instance, Bunge represents the comprehensive version. Jarrett Leplin represents the restrained version. However, sometimes the boundaries between the comprehensive version and the restricted version are diffuse. For instance, although some authors write about SR and identify themselves with SR, other authors define the former as *pragmatic*. Occasionally, critical realists' arguments overlap with restrained version of SR's arguments.

The strong version of SR includes the idea that values can be analyzed by science. This assumption, in turn, follows from the thesis of axiological realism: There are objective values such as health, peace, or equity. The strong version of SR (also known as *integral realism*) supports the thesis of moral realism: There are moral facts, such as unselfish versus selfish actions, and there are true moral principles, such as that unselfish actions promote peaceful coexistence. The restrained version of SR rejects both axiological realism and moral realism.

Naive realism is another kind of realism. In a different way than SR, naive realism claims that the senses provide us with direct awareness of the external world. Naive realism can be equated to common sense's realism: Perception exemplifies unmediated contact with the external world; examples of indirect perception might be seeing a photograph or hearing a recorded voice. Two classic versions of naive realism are the theory of knowledge of reflection and figurative language theory. Naive realists often argue that the multiple neurophysical processes by which we are aware of objects do not entail indirect perception. These processes merely establish the complex path by which direct awareness of the world arrives. The differences in the theory of knowledge between naive realism and SR seem to come from the fact that the first kind of realism does not recognize the crucial distinction, which the second kind of realism introduces, between primary properties (*noumena*) and secondary properties (*phenomena*). This ontological difference produces a number of epistemological differences between the two varieties of realism.

In terms of the theory of testing, SR also conflicts with naive realism's idea that theories are tested more or less like fertilizers. Consequently, for naive realists, the whole methodology of science would be reducible to statistics.

There is a much broader debate between SR and the third kind of realism, namely, *critical realism*. Critical realism constitutes a complex combination of two philosophical traditions, realism and constructivism, that emerges primarily with the oeuvre of Roy Bhaskar (1944–2014) and Wendy Olsen (1960–). For critical realists, natural and social phenomena do not exhaust the category of what really exists in the world. In the work of Bhaskar, there is an ontological distinction between what he calls the domains of the empirical, the actual, and the real. These ontological domains are recognized by most of the critical realists, who conceive of them as their own ontology of the natural and social world. Each of the three domains is relatively sovereign from the other two. The first two can be distinguished as follows: The domain of *the empirical* is made up of human sensory experiences and perceptions, while *the actual* refers to events taking place in the physical world. Perception can sometimes be misleading or unreliable. The third of Bhaskar's domains is *the real*. The real consists of those mechanisms and structures that have causal powers and whose generative capacity creates the order we see in observable entities. Lastly, the real is not the same as the empirical. The empirical gives us an avenue of entrée to the real, but only when the former is guided by theory. Empiricism makes mention only of experience (the empirical), but that events go far beyond what is experienced and that the domain of the *real structures* and other generative mechanisms requires a larger conceptual map of reality. Therefore, empiricism rests upon a deep confusion in terms of the critical realists. It conflates the *empirical* with the *real*, instead of separating them out carefully. Briefly, for the critical realist, the aim of science is the theoretical identification of things and their causal powers.

The literature of critical realism on ontology, epistemology, and methodology is huge, although it has been much more developed in the social sciences. Critical realism suggests a set of philosophical foundations for social research. Taken as a

whole, the methodological bases of critical realism have much in common with strong social constructivism but would strongly modify its underlying assumptions. On the other hand, critical realism has little in common with empiricism. Critical realists have presented a whole series of contributions to methodology. Among these, the most important are the focus on *ontic depth* and the proposal of a new logic of inference, namely, *retroduction*.

Ontic depth refers to having a conceptual map of the world's nature that allows for multiple layers, complexity, interweaving, and dynamic interaction of the sections of that world. *Retroduction* can be defined as the logic of inquiry espoused by critical realists and is the method used to move backward from an observation to the causes of that observation. Retroduction can be distinguished from other research strategies such as deduction or induction. Deduction develops specific claims from general premises. Induction develops general claims from specific premises. As a third mode of inference, retroduction explains events by postulating and identifying mechanisms that are capable of producing them. Strictly speaking, retroduction is not so much a formalized logic of inference as a thought operation that moves between knowledge of one thing to another, for example, from empirical phenomena expressed as events to their causes.

Critical realists suggest the importance of an ontic depth when they, for instance, urge the reality of social structures in the world. By carefully defining structures as real but difficult to observe, critical realists argue that it is possible to carry out many post-structuralist research practices without becoming committed to a markedly relativist or entirely constructivist ontology.

For instance, critical realism suggests that the concept of structure can have two different meanings. It can refer to an ensemble of various things whose relationships produce a single overall entity which *has* a structure. Examples include social structures such as an organization or a city, university as a social institution, the social norm's structure, or the class structure. A second conceptualization of structure emphasizes that the whole has *emergent properties* that differ from the properties of the objects contained by the structure.

The concept of *emergent properties* is very important for critical realists (and also for scientific realists). Emergent properties of a whole are varying over time, if it is a social structure (e.g., social values), whereas they might be somewhat constant over time within a specific physical structure (e.g., the planetary system). Another well-known example of an emergent property arises from the study of the brain: The mind is an emergent property of a brain. Critical (and scientific) realists propose that an ontological property of the things is emergence.

There are several important differences between critical realism and SR. A first difference is epistemological: For critical realists, methods and theory are not value-neutral (e.g., quantitative research is value-laden). A second difference is ontological: For critical realists, empirical regularities do not necessarily point out causality, and the absence of empirical regularities does not necessarily indicate the nonexistence of causality. Therefore, empirical regularities aren't a necessary condition to indicate causality.

Other difference emerges from the use of the term *the real*: For critical realists, *the real* refers to immaterial and impermanent objects which, *for a caused event* or *for a specific observer at a point in time*, are causal and very important. For instance, critical realists asserted that the reality of textually grounded objects was to be checked via the analysis of what social outcomes appeared to be caused by them. A social object (such as a narrative) that has effects can be seen as real. In SR terms, this would mean that narrative is conceived as an independent variable. At the same time, critical realists see narratives not as merely textual but as real social products of ongoing underlying social relations. In SR terms, this would mean that narrative is conceived as a dependent variable.

Philosophy of Science in SR

Contrary to widespread opinion, SR does not claim that our knowledge is accurate. It is sufficient that knowledge be partially true and that some of the falsehoods of our knowledge can be discovered and eventually corrected. Thus, SR postulates the thesis of fallibilism and the thesis of meliorism as two features inherent to science: All knowledge of facts is incomplete and fallible and

much of it is indirect (fallibilism); in principle, any approximation is perfectible (meliorism). Therefore, verifications with reality (the set of empirical tests) will demonstrate again and again that even more precise theories are, at best, more or less close approximations that can be improved.

SR asserts there are two types of properties: primary properties (or noumena) and secondary properties (or phenomena). In contrast, phenomenologists do not admit noumena, such as things or events that exist independently of perceptions (ontological phenomenalism) or elements that, although it considers that they exist, cannot be perceived (epistemological phenomenalism).

Scientific realists do not deny the existence of secondary properties (or *qualia*). In addition, SR postulates that qualia are very different from physical objects and their elementary properties. For instance, SR asserts that secondary properties emerge within a nervous system as a result of external stimuli. In consequence, scientific realists propose that qualia should be studied by psychology, psychophysics, and cognitive neuroscience.

The combination of ontological realism, epistemological realism, semantic realism, and methodological realism suggests a realistic axiology assessment that considers it as a process that occurs in a brain immersed in a social network. In turn, an axiology of this nature invites a realistic interpretation of social norms and rules to address social problems. Likewise, a realistic ethics suggests a practical philosophy that will help address the practical problems both efficiently and in accordance with moral principles. In order to be clear about what SR amounts to in the context of the philosophy of science, it is useful to understand it in terms of those three dimensions: an ontological dimension, an epistemological dimension, and a semantic dimension.

Ontological Theses in SR

Realistic ontology is the basis for an epistemology that encourages the study of real-world properties distinguishing primary properties from secondary properties. In turn, this stimulates a realistic epistemology that begins with a semantic theory of reference and factual truth. Semantic realism encourages such a methodology based on the scientific method. Therefore, it supports the

subject–object dichotomy and encourages the explanation of secondary properties through primary properties.

Ontological realism is the thesis that the universe, or reality, exists *in se et per se*, in and by itself. Ontologically, SR is committed to the mind-independent existence of the world investigated by science. This idea is best clarified in contrast with positions that deny it. For instance, the mind-independent existence of the world is denied by any position that falls under the traditional heading of *idealism*, including some forms of phenomenology, according to which there is no world external to the mind and, thus, independent of the mind. More common rejections of mind-independence stem from neo-Kantian views of the nature of scientific knowledge, which deny that the world of our experience is mind-independent, even if (in some cases) these positions accept that the world in itself does not depend on the existence of minds.

It is important to note that for SR's approach, human convention in scientific taxonomy is compatible with mind-independence. SR accepts the idea of a mind-independent natural kind structure of the world, at the same time that it also accepts that mind-independent properties are often conventionally grouped into a typology by scientific theories, that is, theoretical terms *transform* the external world to the cognitive subject.

Only a tiny part of reality has emerged, remains, and changes for us, but its existence does not depend on the knowing subject. The knowing subject is a real thing surrounded by other real things, most of which existed before the subject, and they did not need their help to exist. The SR proposes four grounds in support of its ontological realism: the arguments of physics, biology, cognitive neuroscience, and history.

The first argument is that fundamental physical laws are invariant under certain changes in frameworks, including observers. Frameworks are special physical systems that do not necessarily include humans. Biology's claim asserts that organisms, including bacteria and philosophers, extract nutrients and energy from their environment and they are provided with sensors for external signals. This is the reason why organisms die if they are completely isolated from their environment, and they face great risks when their sensors

perform poorly. The brain cannot exist without the world around it, from which it extracts not only nutrients but also the stimulation necessary for its normal development and functioning. The role of the environment in the development of a normal brain has been shown experimentally. The genome is not sufficient for the normal development of the brain: Stimulation from the external world is also needed. In short, environmental stimuli contribute heavily to the development of the brain. When these stimuli are absent, the brain fails to develop normally.

The argument of neuroscience can be summarized as follows: Mental functions are brain processes, some of which represent external events. Scent comes from an external signal. Sometimes the brain produces illusions or hallucinations: It sees *water in the desert* or it hears *voices*. On the other hand, no two people see the world exactly the same way, as no two brains have identical life stories. However, all people, even those who have severe autism, agree that there are certain things that exist by themselves, such as trees, buildings, and others.

Finally, the argument of history points out that all historical sciences, from geology to cosmology, to evolutionary biology, and to historiography, take the past for granted. Moreover, they assume that the study of the past cannot change the past. Constructivists could say that the past (not only their study) is a social construction. For SR that affirmation shows confusion between the past, history and its descriptions, and historiography.

Epistemological Theses in SR

Epistemologically, SR is committed to the idea that theoretical claims (interpreted literally as describing a mind-independent reality) constitute knowledge of the world. This idea contrasts with skeptical positions which, even if they grant the ontological and semantic dimensions of realism, doubt that scientific investigation is epistemologically powerful enough to yield such knowledge or, as in the case of some antirealist positions, insist that it is only powerful enough to yield knowledge regarding observables.

Epistemological realism can be synthesized, according to Mario Bunge, by one definition and six theses. The definition is this: If a thing is real,

this thing exists independently of the cognitive subject. Note that this definition cannot be a criterion of reality, since it does not tell us how to find out if that thing persists while we are not looking it. The six theses are as follows: (1) Reality is understandable, or better, you may know some facts, but usually in a partial and gradual manner; (2) indirect knowledge is deeper and is obtained by means of theories and indicators rather than through mere perception or intuition instantaneous; (3) almost certainly, sometimes we err, for instance, by using poor operational techniques, making false assumptions, or producing false data; (4) it is possible, in principle, to improve any given piece of knowledge, and make it more accurate or deep; (5) moderate pluralism: In principle, any set of facts can be presented by means of alternative hypotheses or theories; and (6) objective knowledge supported by solid tests and valid theories is superior to subjective intuition.

SR supports the so-called *explorer's argument*: Anyone who intends to explore a territory or investigate a thing, event, or specific process assumes that these things exist or may exist. Thus, the physicist who tries to find Higgs bosons with the aid of particle accelerators of high energy supposes the possibility that these particles, postulated by a theory, exist.

Semantic Theses in SR

Semantically, SR is committed to a literal interpretation of scientific claims about the world. According to SR, claims about scientific entities, processes, properties, and relations, whether they are observable or unobservable, should be construed literally as having truth-values, whether true or false. This semantic commitment contrasts primarily with those of so-called *instrumentalist epistemologies of science*, which interpret descriptions of the unobservable simply as instruments for the prediction of observable phenomena.

Semantic realism is the concept according to which: (1) Some propositions refer to (talk about) facts, and (2) some factual propositions are true to some extent. Thesis (1) is rejected by the textualists and deconstructionists, according to whom a symbol refers to other symbols. Radical skeptics reject thesis (2). Note the difference that semantic realism raises between reference and

representation. While tautologies, conventions, and statements of pure mathematics have referents, they do not represent anything in the real world.

Differing from the redundancy theory (the concept of truth is redundant since to state p is the same as claiming the truth of p), from the coherence theory (a proposition is true in a body of propositions if and only if it is consistent with the remaining constituents of this body), and from the correspondence theory (a factual proposition is true if and only if corresponds to [represents] the facts to which it refers), SR postulates the synthetic theory of truth.

The synthetic theory of truth can be summarized by a definition and a criterion of factual truth. The definition is this: Proposition p asserting fact b is true if and only if f occurs. The criterion of factual truth includes: (1) p is compatible with (and, exceptionally, equivalent to) the pertinent empirical evidence, and (2) p is consistent with the bulk of the relevant background knowledge. Note the distinction between the definition and the criterion of factual truth: the first tells us what is the truth and the second tells us how to identify it. The failure to distinguish between truth and evidence of truth is the source of some varieties of antirealism. One of them is the pragmatic assertion that there are no bridges between our theories and the world, because these theories can only describe the relevant evidence. According to SR, such confusion between reference and test is like confusing the bacteria with the microscope. This confusion also characterizes the verification theory of meaning, proposed by the Vienna Circle. According to SR, this theory is false: Measurements and experiments can tell us whether a hypothesis is true, not what this hypothesis means. Moreover, the meaning precedes the testing, because we need to know what a hypothesis means *before* designing proof of its truth value.

The *error argument* is another of the postulates of semantic realism. The very concept of scientific error, either conceptual or empirical, presupposes the real existence of the postulated entity. For example, a scholar can be interested in the number of crimes in a city. Then, the researcher seeks the reports on crimes in the police records. However, this number of crimes that the researcher believes

that there were in the city is wrong because many victims of such crimes, especially in countries where the police have poor performance (if they are not openly corrupt), do not report the crime to the police. There are victims who have not been included in the scholar's account. Consequently, those schools of philosophy that reject the concept of objective truth cannot use the concept of error understood as discrepancy between theory and reality.

Methodological Theses in SR

In the methodological dimension, SR holds two main theses: (1) The scientific method is the best approach to get the most objective, more accurate, and deeper truths about events of all kinds (methodological realism) and (2) this postulate applies to both natural science and social science (methodological monism). Postulate (1) is the more controversial in the debate between realists and antirealists. Postulate (2) is the most contested issue between SR and critical realism.

Opponents of methodological monism argue that subjectivity cannot be studied scientifically. They affirm that cultural or social sciences, unlike the natural sciences, must seek to *understand* (*Verstehen*) social facts rather than *explain* them (interpretivist thesis). To understand social facts, interpretivism's method proposes to find the *meaning* that actors give to their behavior. Natural facts, unlike social facts, are not guided by meaning. The different nature of social facts and natural facts must sustain an insurmountable methodological dualism between social sciences and natural sciences.

Methodological realism recognizes the differences between natural sciences and social sciences regarding their objects of study and the techniques to study them. However, methodological realism also argues that such differences do not represent any obstacle to the use of a general method in all sciences. To justify the latter assumption, methodological realism postulates that everything that exists encompasses a material dimension, including the meaning that individuals give to their behavior. Such meanings are manifested in processes occurring in the prefrontal cortex that are studied by cognitive neuroscience. Finally, methodological realism asserts that all researchers in all

sciences (natural and social) have similar brains, and all researchers in all sciences (natural and social) share the goal of finding objective truths.

Rodolfo Sarsfield

See also Constructive Empiricism; Instrumentalism; Phenomenalism; Semantics (Scientific and Empirical)

Further Readings

- Bhaskar, R. A. (2008/1975). *A realist theory of science*. Routledge.
- Bunge, M. (1993). Realism and antirealism in social science. *Theory and Decision*, 35(3).
- Bunge, M. (1998). *Philosophy of science*. Transaction, 2.
- Bunge, M. (2003). *Emergence and convergence: Qualitative novelty and the unity of knowledge*. University of Toronto Press, 2003.
- Bunge, M. (2006). *Chasing reality: Strife over realism*. University of Toronto Press.
- Laudan, L. (1990). *Science and relativism*. University of Chicago Press.
- Leplin, J. (1984). Introduction. In J. Leplin (Ed.), *Scientific realism*. University of California Press.
- Leplin, J. (1997). *A novel defense of scientific realism*. Oxford University Press.
- Olsen, W. (Ed.). (2010). *Realist methodology* (4 vols. SAGE Benchmarks in Social Research Methods). SAGE Publications.
- Popper, K. R. (1959/1935). *The logic of scientific discovery*. Heinemann.
- Popper, K. R. (1963). *Conjectures and refutations: The growth of scientific knowledge*. Basic Books.
- Popper, K. R. (1972). *Objective knowledge: An evolutionary approach*. Oxford University Press.
- Putnam, H. (1982). Three kinds of scientific realism. *Philosophical Quarterly*, 32(128), 195–200.
- Putnam, H. (1984). What is Realism? In J. Leplin (Ed.), *Scientific realism*. University of California Press.

SCIENTIFIC REVOLUTIONS

Scientific knowledge typically accumulates over time and transforms gradually over the course of years, as new studies and observations provide fresh insights into existing problems and ideas. Every so often, however, a powerful research finding or radically different proposition challenges

the foundations or assumptions that characterize the scholarly pursuits of the time, rendering some theories obsolete and opening up new pathways for inquiry. When this happens, it is referred to as a *scientific revolution*. Therefore, a scientific revolution constitutes a radical change in the conceptualizations that are accepted within one or more disciplines. In order to account for the fact that revolutionary ideas may also emerge in fields that are not considered strictly scientific, the term *conceptual revolution* is used in many contemporary contexts.

Changes in the knowledge base of a learned community occur normally over time, but what defines a revolution is a profound shift in perspective leading to new ways of thinking about old problems, new problems being defined as appropriate for study, or new processes being laid out for the pursuit of knowledge. According to philosophers of science, revolutions seem to follow specific patterns, are usually preceded by a period of crisis and uncertainty, and can be sparked by the work of people who are new to established intellectual traditions. Revolutionary ideas can stem from a variety of fields and often end up influencing the theoretical constructs used in more than one discipline.

In historical terms, we refer to the 16th and 17th centuries as encompassing a particular kind of scientific revolution—one that defined the modes of inquiry and the kinds of questions that would occupy scholars in subsequent centuries, directly shaping the scientific and industrial development of contemporary societies. It was during this time that the scientific method was formalized, the principles of rationalism were established, and the Enlightenment project emerged as a pathway for achieving progress, with special emphasis being placed on education, reason, and empirical research. Thus, the significance of the Scientific Revolution resided in its ability to impact not only scholarly practices but also people's common understandings about the world, the cosmos, and humanity's place in it. In later centuries, deep shifts in theorization and inquiry have continued to take place, accompanying new breakthroughs in science and research, and yielding new insights into a multitude of problems. The following sections discuss the process leading to scientific revolutions; the factors contributing to

the emergence of modernity and the scientific revolution of the 17th and 18th centuries in Europe; the nature and applications of the scientific method; the development of rationalism, empiricism, and the philosophy of science; and the 20th century's conceptual revolutions in the sciences, including the emergence of qualitative methods, critical theory, and the development of new disciplines.

The Emergence of a Scientific Revolution

A revolution is often brought about by a discovery or proposition that forces a scrutiny of previous assumptions: This may be prompted by staggering research findings, new technologies, or conceptual reformulations that cannot be disregarded. In many cases, a scientific revolution involves a breakthrough that transcends the scientific community. Among the prominent scholars who have written about the process leading to revolutions is Thomas Kuhn (1922–1996). His book *The Structure of Scientific Revolutions* outlines how these revolutions come about, the possible paths that they follow, who are the people most likely to initiate them, and their importance for scholarly progress.

Kuhn argues that revolutions may come from two main sources: (1) a major breakthrough or empirical finding or (2) a new line of thought that gains momentum and is incompatible with previously accepted conceptualizations. In this sense, a scientific revolution brings about theoretical reformulations and a necessary reassessment of previously established truths. For Kuhn, theoretical and epistemological constructs have an orienting role that shapes not only the kinds of explanations researchers arrive at when considering a phenomenon but also the very questions they ask and the observations they conduct. He describes a *paradigm* as a broadly influential framework that shapes more specific research programs and theoretical constructs, within and across various disciplines. When accepted paradigms begin to lose their explanatory power, the opportunity arises for the emergence of new ones that will eventually replace them, leading to a shift in perspective. The concept of paradigm is connected to that of *disciplinary matrix*, which encompasses both the problems and the investigative practices that are normatively addressed within a discipline.

Kuhn proposes that revolutions may evolve in different ways: For example, radically new ideas may be marginalized for a while or quickly brought to center stage, openly debated, and tested. They may spark deep disagreement between different members of a scholarly community, or they may be embraced by influential individuals who help validate and promote them. Interestingly, the people most likely to come up with a revolutionary idea are often those who are not already situated high in the ranks of academic power: They may be very young or not fully schooled in the established ways of a particular discipline. Because they have not yet endorsed existing paradigms, these individuals are able to bring a fresh perspective to old problems.

Scientific revolutions often take place over a period of time, and they can occur in all disciplines. They always create some degree of conflict and uncertainty, as the shift from an established set of conceptualizations and expectations to a newer one moves through a period of transition and redefinition. Nevertheless, once the new ideas become broadly accepted, they are incorporated into the established ways of thinking and doing of the scholarly community, therefore losing their revolutionary character. So, the larger cycle begins again, as any established conceptualization is likely to be eventually challenged by new research or new theoretical constructs.

Another prominent philosopher of science, who influenced Kuhn's ideas on theory development, is Karl Popper (1902–1994). One of the contributions made by Popper involves his exploration of the characteristics that define a scientific theory as such, separating it from pseudoscientific or folk knowledge. In his book *Conjectures and Refutations*, Popper discusses the main ways that theory developments occur, bringing along with them scientific or conceptual revolutions. Popper argues that theories can be either supported or disproven, so research that attempts to test an existing theory leads to either a confirmation of its basic propositions or a refutation that forces its reassessment. However, many theories are so big, abstract, and complex that they can hardly be *proven* in definitive terms as positive explanations of phenomena; yet as long as they cannot be *proved wrong*—or falsified, as he says—they continue to stand as plausible explanations. Particular

aspects of them may be tested, and the evidence may either support the theory's propositions or not, thus forcing an examination of the whole that can lead to new ideas. While discussing this, Popper engages with issues of epistemology, also considering topics such as validity and normalization within scholarly debates—which are also connected to theory change.

Overall, scientific revolutions, insofar as they represent a paradigm shift and the emergence of new disciplinary matrices, may have an impact that stretches across a variety of fields. Because this may include the humanities and other areas not considered strictly *scientific*, authors such as Paul Thagard have proposed a broader term: *conceptual revolutions*, which designates a profound shift in perspective that has theoretical and practical implications. Revolutions facilitate the introduction of new theoretical constructs, which may compete for acceptance with preexisting theories. Some fields may be less likely to tolerate competing conceptualizations without allowing for the emergence of a dominant framework. However, in other fields, a variety of theories may coexist as alternative explanations for the same phenomena, leading to parallel conceptual definitions and research agendas.

Historical Changes and Epistemological Shifts During the Renaissance

Popper proposes that a theory can be seen as a tool or lens that scholars use in order to explain and understand what happens in the world—and this tool or lens is the product of its time and place. Theories may be proposed by individual authors, but it is their broad acceptance by an academic community that determines the extent of their influence within a scholarly field. Regardless of attempts to establish science as an objective and neutral realm, research is inscribed within historical contexts. The scientific revolution was the result of profound intellectual changes that took place in Europe during the 16th, 17th, and 18th centuries. This period enabled the development of all the natural and social sciences and influenced a number of social, economic, and political transformations that later shaped the modern world. For example, the rise of urbanism

and industrialization, the founding of multiple institutions of higher education, and the allocation of state funds to support scientific pursuits accompanied the emergence of new health programs, technological advancements, and bureaucratic processes.

The conceptual shifts that shaped the scientific revolution in Europe emerged during the period that spans from the Renaissance to the Enlightenment. During the preceding period, the Middle Ages, dominant views of nature, humanity, and life were shaped by the principles of the Catholic faith, so the goal of all studies was to discover truths that honored God and his creation. This implied a particular epistemological framework that also limited the kinds of phenomena deemed suitable for inquiry: The pursuit of other kinds of knowledge (e.g., about the biological structure of the human body) was strongly discouraged, and propositions that challenged the established order were harshly sanctioned. Institutions such as the Inquisition exercised a disciplining role in numerous territories. Nonetheless, the advent of the Renaissance brought along a radical reformulation of dominant ideas regarding nature, society, humanity, and knowledge, both within intellectual circles and larger communities. Many of these ideas were underpinned by a shift of emphasis from the divine to the natural world, reigniting an engagement with worldly issues, as religious matters lost their privileged position and were replaced by an interest in emergent political movements. This shift in perspective brought along a reassessment of humankind's place within the natural order and yielded new conceptual frameworks throughout the sciences.

The Renaissance began in what is now Italy and involved a rediscovery of writings by classical philosophers, poets, astronomers, and mathematicians of classical antiquity. Some of them—such as Plato and Aristotle—were already studied at the time, and their ideas were incorporated into canonical scholarly works. However, many others had been largely ignored since the fall of the Western Roman Empire in the 5th century C.E., even as numerous manuscripts were preserved in the libraries of the Eastern part of the empire, also known as *Byzantium*. After the fall of Byzantium in the 15th century, several factors facilitated a large conceptual shift among learned communities

in Europe: Classical works were rediscovered, inventions such as the printing press made books available to a wider population, and observations conducted by scholars such as Nicolas Copernicus and Galileo Galilei challenged the dogmatic ideas that had prevailed until then. The establishment of the first universities, starting in the 11th and 12th centuries, had helped formalize areas of study such as theology, philosophy, medicine, and the law, and the expansion of educational institutions in the centuries that followed shaped the ensuing intellectual debates, creating spaces for new theoretical understandings.

The beginning of the Renaissance was marked by major political crises and significant social changes. Former feudal kingdoms merged and gave rise to new nation-states that claimed specific territories and soon initiated projects of conquest and expansion. Challenges to medieval ways of life and modes of thought included a critique of the serfdom-based loyalties that supported feudal institutions and a reshaping of the social structure in order to accommodate an emerging class of traders and bankers. As Christopher Columbus and other adventurers traveled to the New World, evidence of new geographical territories forced a reassessment of old cartographic knowledge and a redrawing of maps. The perceived mystery of the Americas soon heralded a new era of exploration and colonization that brought both tangible and conceptual novelties to the European continent. The conquest effort, marked by violence, would lead to centuries of clashes with indigenous communities and the enslavement and domination of specific groups of people, which further shaped social and political affairs around the world. Meanwhile, the religious reformation that began in Europe in the 16th century and led to the emergence of Protestantism sparked a wave of religious conflict, bringing about several lengthy and violent wars that changed the landscape of the continent.

In the midst of all this, a profound change in perspective took place within scholarly communities, enabling paths of inquiry that were not previously available. Renaissance scholarship encouraged an active engagement with multiple arts and sciences and promoted curiosity and experimentation. Two of the most iconic figures of the Renaissance are Leonardo Da Vinci and

Erasmus of Rotterdam, both active in the 15th and 16th centuries. Leonardo's interests and extraordinary skills ranged from the arts to engineering, botany, and writing, and he was among the first to venture into investigating previously forbidden matters such as the inner workings of the human body, producing anatomical drawings that were used in the study of medicine. Erasmus of Rotterdam, a Dutch humanist and scholar, revolutionized philosophy by placing humanity at the center of his inquiry, instead of divine and spiritual affairs. In broader terms, Renaissance thought questioned generally held assumptions regarding humanity, nature, good, evil, and God, bringing along the possibility of engaging with many new topics. The changes that occurred during this period had far-reaching consequences from the point of view of epistemology and theorization, and established the foundation for the emergence of new disciplines and new scholarly practices.

The Scientific Revolution and the Scientific Method

The Scientific Revolution brought along a way of thinking that disregarded folk wisdom as a source of knowledge, putting emphasis on direct observation and evidence in order to draw general conclusions. It encouraged a detailed recording of the logical steps followed during the inquiry process and facilitated the accumulation and dissemination of knowledge through written treatises and manuals. This marked an epistemological and methodological shift with regard to the practices that were used previously, when explanations of natural or historical events were often codified in myths and folklore that were transmitted through oral or literary traditions. The emphasis on direct observation and measurement also challenged the abstract philosophical deductions used in many scholarly pursuits, such as Aristotelian syllogisms.

One of the best examples of a paradigm change that shook widely held assumptions and had an influence across the disciplines during the time of the Scientific Revolution was heliocentrism, which was first proposed by Nicolas Copernicus and later defended by Galileo Galilei. An Earth-centered conceptualization of the planetary system

had prevailed for more than a thousand years, so the idea of a solar system in which the planets circled around the sun was groundbreaking. This had implications not just for the way people studied the cosmos but also for the way they thought about their own place in the universe; furthermore, because the new model challenged the perceived order of things, it was deemed politically dangerous. When Copernicus's work was published in 1543, it was controversial. The evidence later gathered by Galileo, a 17th century mathematician, raised the matter's profile even more, and the Inquisition stepped in to suppress it. It was only after Galileo's death that the accuracy of his observations was determined, leading to a paradigm shift that opened new pathways for inquiry into the workings of the universe.

The Scientific Revolution led to the establishment of a range of new disciplines, including biology, chemistry, and engineering. Significantly, it produced a particular protocol for conducting research—the *scientific method*—which continues to be used today. The method provides a step-by-step procedure that scholars can follow in order to arrive at valid results, and its guiding principles are objectivity, controlled manipulation of variables, and conceptual separation between the observer and the observed. The purpose of research conducted with the scientific method is usually to predict future outcomes and accrue knowledge that can potentially have practical applications and solve specific problems. Moreover, the scientific method establishes normative guidelines for conducting and evaluating research in a manner that is systematic and organized. The basic steps of the method are four: question, hypothesis, test, and analysis. Thus, the process begins with a preliminary observation of phenomena and the drafting of research questions. This is followed by a review of relevant literature and the formulation of one or more plausible hypotheses. The next steps, test and analysis, involve designing an experiment or conducting systematic observations in order to collect empirical proofs. Testing is ideally conducted within a controlled environment, which requires isolating and manipulating variables so as to observe specific outcomes. The analysis of the results is to be done objectively, with the help of inferential statistics and mathematical formulas, so this kind of scholarly work is

also called *quantitative research*. Finally, the goal of the analysis is not just to observe general trends but to discover statistically significant relationships that can be affected by induced manipulations and are unlikely to occur by chance.

The scientific method is connected to a particular epistemological stance: namely the idea that the workings of the world can be objectively uncovered through direct observation and measurement and that future outcomes can be predicted on the basis of recorded data. The protocols that accompany the scientific method also establish specific processes for the development, refutation, or proving of new theories, which are supposed to arise in connection with a research question and must represent an effort to account for accrued evidence. Insofar as any research project that uses the scientific method is based on a review of literature and new findings can lead to new research projects, the scientific method points to a steady accumulation of knowledge. Furthermore, the results of research done with the scientific method are expected to be consistent and replicable—meaning that similarly conducted observations should yield similar findings—which shapes how new discoveries are handled by scientific communities. Because of all this, the scientific method establishes a particular way of examining the world and producing knowledge. Today, the method is used in all the biological and physical sciences, as well as many of the social sciences.

Philosophy of Science, Rationalism, and Enlightenment

In the 18th and 19th centuries, thinkers such as Immanuel Kant (1724–1804) and G.W. F. Hegel (1770–1831) strived to articulate a *philosophy of science* that would engage with debates on ethics, reasoning, and human experience, while also considering humanity's place within the natural world. Kant's *Critique of Pure Reason* focuses on issues of epistemology and validity in logical inquiry, while Hegel's writings emphasize the importance of systematic analysis and advance a teleological view of history that greatly influenced later writings on dialectics and materialism. The emphasis on reason that appears in the work of both philosophers is connected to the rise of an intellectual current that shaped epistemological and

theoretical debates within scientific fields: rationalism. Both Hegel's ideas regarding the primacy of "pure Thought" and René Descartes's (1596–1650) often-quoted maxim "I think, therefore I am" constitute examples of rationalist philosophy.

The rationalist view of humans as intellectual beings who could dominate and rule a mindless natural world was far removed from medieval conceptualizations of humanity and nature that were centered on the acceptance of a divine order. Rationalism conceived the human intellect as the primary force of history, validating attempts to manipulate the natural world by means of technology and industry. In tune with the perspectives that reigned during the Scientific Revolution, rationalism valued logical reasoning, a careful collection of data, systematic analysis, and detailed record-keeping, in order to arrive at a conclusion, truth, or prediction. Its proponents often engaged with political and economic issues of their time, approaching problems with a mindset that focused on the maximization of resources. This epistemological approach was eminently represented in the ideas of authors such as John Locke (1632–1704), Francis Bacon (1561–1626), and Isaac Newton (1643–1727), whose work would range from propositions regarding government to advancements in engineering.

The intellectual landscape of Europe during the Scientific Revolution also enabled the development of a line of thought that later gave form to a political and social project: the Enlightenment. Enlightenment ideals embraced the principles of rationalism, but proponents added a specific concern with progress and strived toward the betterment of the human condition, paying particular attention to such aspects as education and economic development. Their connection to science resided in the belief that technological and scientific advancements would provide the means for solving all kinds of problems—a belief that was underscored by an optimistic view of human nature and societal progress. Enlightenment thought engaged with revolutionary ideas such as universal education, industrialization, market economics, social reform, and representative democracy, giving rise to new theories of government, society, and human behavior. Within the political sphere, the epistemology of the Enlightenment

shaped projects that influenced historical events such as the French Revolution and the emancipation of the American colonies. Some of its major proponents were Voltaire, Montesquieu, Jean-Jacques Rousseau, David Hume, and Adam Smith, whose ideas influenced generations of scholars.

Enlightenment thought provided the backdrop and the fuel for the extraordinary growth that the sciences experienced during the 19th century and the beginning of the 20th century. At that time, radically new conceptual propositions—such as Charles Darwin's *theory of evolution* and Albert Einstein's *theory of relativity*—were developed, which led to theoretical reformulations and a wealth of empirical research in various fields. Although some aspects of the Enlightenment project were later critiqued, many of its fundamental principles continue to hold appeal, especially regarding the relationship between reason and research, political and social progress, and science and development.

20th-Century Conceptual Revolutions

The 20th century saw both an extraordinary growth in the theoretical base of the natural and social sciences, and the emergence of a critique of the principles orienting research and theorization since the Scientific Revolution. The Enlightenment's trust in the promise of science and technology as a means for solving human problems contributed to the prominent place that the sciences were given within the scholarly community throughout the 19th and 20th centuries. Breakthroughs in health care, such as the discovery of vaccines and antibiotics, and new developments in transportation, telecommunications, and space exploration gave credence to the belief that applied science could hold the answers to many issues, directly impacting people's quality of life in modern societies.

The epistemological perspective that defined much of the scholarly work done during those two centuries was empirically oriented. Empirical research used the scientific method and embraced the premise that the world is objective, the answers to any questions about the world can be found by well-crafted research designs, and good science is neutral in terms of its processes and consequences. Empirical approaches were

dominant within the research carried out within new disciplines such as psychology and sociology, in addition to the hard sciences, and shaped the theoretical constructs that emerged from those fields. For example, *behaviorism*—a theoretical framework that used empirical and positivistic methods to investigate living organisms—was proposed by the psychologist B. F. Skinner (1904–1990) and emphasized observable behavior and cause-and-effect reactions, becoming influential within various disciplines, including psychology and education.

World affairs that defined the 20th century also impacted scholarly communities and led to new conceptual formulation. For example, the two world wars reshaped the political map of Europe and global power dynamics; increased urbanization and industrialization changed the character of city life and the nature of labor; and technological innovation changed the way information was produced and disseminated throughout the world. This was accompanied by significant socioeconomic and cultural shifts that also became the topic of academic debate. On a different level, scholarly practices were further shaped by factors such as the creation of many new public universities, the introduction of corporate funding into some research arenas, and the creation of worldwide academic networks and professional associations.

In an effort to develop epistemological and theoretical tools to understand the social and political transformations going on at the time, new approaches to research were developed, some of which constituted conceptual revolutions in their own right. Critics of empirical research argued that the framework had limited explanatory power for problems involving naturalistic and highly complex social environments, where observable variables and measurements could not account for multiple and simultaneous interconnections. Hence, during the second half of the 20th century, two emergent epistemological and conceptual frameworks had a significant impact across many fields: *cognitivism* and *critical theory*. They both challenged assumptions derived from rationalism and the Enlightenment, and they both were foundational for the development of new disciplines within the academy.

An interest in the workings of the human mind greatly increased throughout the 20th century—science had made enormous progress in its exploration of the natural world, but the human brain remained largely a mystery. The introduction of *psychoanalysis* as a theoretical framework had shaken many social norms, with Sigmund Freud's propositions regarding the unconscious, the ego, and innate sexual drives impacting various intellectual circles and artistic movements. Yet the core ideas of psychoanalysis could hardly be proven through empirical means. Partly aided by the development of new surgical instruments and new devices that could be used to obtain glimpses into the workings of the brain and therefore of the mind, research on the brain and brain disorders showed a marked growth around the 1970s. With it, a new conceptual framework emerged: *cognitive theory*. Whereas psychoanalysis had focused on subconscious drives and the effects of traumatic experiences, and behaviorism had bracketed the *black box* of the mind in order to focus on observable behavior, cognitivism sought to discover the internal logic and the hidden structures of the mind, as well as of human faculties such as thought, memory, and language. Hence, it sought to make sense of both the biological and social aspects that influenced people's thoughts and behaviors.

The epistemological framework that cognitivism brought forth deeply impacted disciplines such as psychology and linguistics, which experienced a cognitive turn that modified not only the problems tackled by standard research projects but also their methods and their goals. For example, linguistic analysis had been carried out for years on the basis of comparative, anthropological, or philological approaches, but cognitivism brought along new questions regarding the structure of language and the mind, which were revolutionary. New thinkers such as Noam Chomsky (1928–) and Steven Pinker (1954–) developed propositions regarding generative linguistics, linguistic universals, and language development. Impacting a number of other disciplines, the cognitive revolution brought with it a direct engagement with questions regarding the connection between the mind, the brain, and the world, stirring up epistemological debates within the humanities and social sciences.

Meanwhile, the social and political changes that occurred in the 20th century led other thinkers to develop a theoretical and methodological framework that directly challenged the principles of the scientific method: *critical theory*. The Nazi regime, the establishment of the Union of Soviet Socialist Republics, the Civil Rights movement, the feminist movement, and other historical events and cultural transformations prompted a scrutiny of the principles underlying the frameworks of analysis that had been in use since the Enlightenment. Critical scholars dismissed the idea that research should be neutral, objective, and separated from political affairs, instead seeking to directly engage with topics such as ideology, power, and inequality in contemporary society. Their goals were to scrutinize structures of domination and contribute to social change, while empowering and reasserting the validity of marginalized voices and experiences. In this sense, critical theory posed an epistemological shift that would yield new methods of analysis, later grouped under the umbrella of *qualitative research*. Because of this, its introduction in academe was revolutionary.

Critical theory developed in connection to a critique of the status quo and a desire to support social transformation. Early landmarks in the formulation of its principles included the publication and dissemination of Karl Marx's *Capital*, which presented a comprehensive critique of capitalism as a system that structures not just economic relations but also social and political formations. The arguments put forth by Marx were taken up by many later authors, and the influence of Marxism went well beyond intellectual circles, leading to the creation of revolutionary political movements. Within the academy, critical theory was enriched in the 1940s and 1950s by the work of scholars who came to be collectively named the *Frankfurt School*—a group that included Walter Benjamin, Theodor Adorno, Max Horkheimer, Herbert Marcuse, and later, Jürgen Habermas. Horkheimer, Adorno, and Marcuse were among the many intellectuals and artists who emigrated to the United States to escape from Nazi Germany. Their writings were influenced by European philosophers such as Hegel, and their analysis of contemporary U.S. American society was foundational for the

emergent theoretical framework. Adorno, for example, was critical of the goals of empirical research and the role that scientific pursuits had in America, where they apparently served the interests of large corporations and the State. He was also concerned with ethical matters, such as the way scientific pursuits could be linked to abuse in the name of a higher good.

In the 1960s and 1970s, as social movements for civil rights, peace, feminism, and gender equality gained prominence, intellectuals who belonged to groups that had been historically marginalized from scholarly communities began to challenge the emancipatory promises of modern science. They argued that science tended to either obliterate or objectify the experience of women and people of color, while masking a myriad of biases under a cloak of neutrality. Furthermore, they stated that theory and research are never objective because dominant ideologies percolate conceptual constructs and methodological protocols. For example, Michel Foucault (1926–1984), author of *The Archaeology of Knowledge*, explains that conceptual revolutions go hand in hand with larger social changes, advocating for a rethinking of contemporary social institutions, which stand in connection to discursive practices and ideological formations.

Critical scholarship influenced the development of other theoretical frameworks, such as cultural studies, which originally developed in England in the 1960s and directly supported interdisciplinary conversations. Cultural studies scholars also rejected quantitative research, arguing that it reduced social phenomena to measurable bits, and instead embraced *grand theory* because they believed it offered better possibilities for explaining the historical complexity of social realities. This led to the development of new objects of research, such as media, and new methods of qualitative inquiry. Qualitative methodologies such as ethnography, text analysis, and structured interviews gained an increased validity during the 1980s and 1990s, and today, they are widely used—their introduction prompted a conceptual revolution insofar it led to a deep reformulation of the standards that were associated with scholarly research. Besides entering fields such as sociology, communication, literature, and political science, critical and cultural theory greatly influenced

the epistemological perspectives used in new disciplines such as Women's and Gender Studies and Ethnic Studies.

In sum, scientific or conceptual revolutions can be brought along by new discoveries or novel reformulations of existing ideas, which force a scrutiny of the standard theoretical perspectives and research methods used within a given discipline. Significant advancements in science and technology may bring about a specific scientific revolution, as they have done throughout history, but in broader terms, a revolution that reshapes the conceptual base of a scholarly field can take place in any discipline and in any period of time. Often, new and revolutionary perspectives do not emerge from the higher ranks of an established field but come from the margins or from neighboring areas of study. Conceptual revolutions may be sparked by major discoveries or unanticipated findings that force a revision of existing theories or by forceful arguments that shed new light on old assumptions. Epistemological and philosophical inquiries regarding the nature of knowledge have significantly influenced scientific and scholarly pursuits during the last centuries—and have been influenced by them as well. The history of the modern world has been shaped by developments in science, and many contemporary standards regarding scholarly research refer back to the protocols developed during the Scientific Revolution. Within the academy, newer conceptual revolutions have brought along new research methodologies and defined new objects of study, which have impacted the scholarly work done across many disciplines.

Claudia Bucciferro

See also Epistemology; Paradigm; Philosophy of Science; Rationality; Social Construction of Scientific Knowledge; Theory Change

Further Readings

- Adorno, T. (2008). *Lectures on negative dialectics*. Polity Press.
- Chomsky, N. (2002). *New horizons in the study of language and mind*. University of Cambridge Press.
- Dear, P. (2001). *Revolutionizing the sciences: European knowledge and its ambitions, 1500–1700*. Princeton University Press.

- Descartes, R. (1935). *Discourse on method* (J. Veith, Trans.). Open Court Publishing.
- Foucault, M. (1972). *The archaeology of knowledge* (A. M. Sheridan, Trans.). Pantheon.
- Geuss, R. (1999). *The idea of a critical theory: Habermas and the Frankfurt school*. Cambridge University Press.
- Grasswick, H. (2011). (Ed.). *Feminist epistemology and philosophy of science: Power in knowledge*. Springer.
- Hall, S., Hobson, D., Lowe, A., & Willis, P. (Eds.). (1996). *Culture, media, language*. Routledge.
- Hoyningen-Huene, P. (2012). Philosophical elements in Thomas Kuhn's historiography of science. *Theoria*, 75, 281–292.
- Kant, I. (2007). *Critique of pure reason* (M. Weigelt, Trans.). Penguin Books.
- Kuhn, T. S. (1962). *The structure of scientific revolutions*. The University of Chicago Press.
- McRae, R. (1961). *The problem of the unity of the sciences: Bacon to Kant*. University of Toronto Press.
- Nnaji, J., & Luján, J. L. (2016). The content of science debate in the historiography of the scientific revolution. *International Studies in the Philosophy of Science*, 30(2), 99–109.
- Popper, K. (1963). *Conjectures and refutations: The growth of scientific knowledge*. Routledge.
- Rosen, M. (1982). *Hegel's dialectic and its criticism*. Cambridge University Press.
- Sperry, R. (1993). The impact and promise of the cognitive revolution. *American Psychologist*, 48(8), 878–885.
- Thagard, P. (1992). *Conceptual revolutions*. Princeton University Press.

SEMANTICS (INTRODUCTION TO THEORY)

Semantics is the part of linguistics that is concerned with meaning. Meaning is a notion with a wide range of applications, some of which belong to the field of semantics while others lie beyond it. Meaning is always the meaning of something. Words have meanings, as do phrases and sentences. Semantics is exclusively concerned with the meanings of linguistic entities such as words, phrases, grammatical forms, and sentences, but not with the meanings of actions or phenomena. The meanings of words and sentences cannot be studied independently of how they are actually

used in speech. After all, it is language use that provides the data for semantics. Therefore, the meanings of linguistic utterances also matter to semantics.

The technical term for using an expression for something is *reference*. The function of the pronoun *I* is reference to the speaker of the sentence: When people use *I*, they refer to themselves. The entity referred to by an expression is called its *referent*. So the *referent* of *I* is always the speaker. The meaning of the pronoun can thus be described as follows: *I* indicates reference to the speaker. Similarly, the pronoun *you* indicates reference to one or more addressees.

This entry describes and explains the various forms of meaning addressed within the domain of semantics and concludes with some general observations about the relationship between the study of semantics and the modern approach to linguistics.

Expression Meaning and Utterance Meaning: Truth Condition

The meanings of words, phrases, and sentences, taken as such in their general sense, apart from any particular context, constitute the level of meaning which will henceforth be called *expression meaning*. *Expression* is just a general term for words, phrases, and sentences. The term *expression meaning* covers, in particular, word meaning and sentence meaning. The level of *expression meaning* constitutes the central subject of linguistic semantics. It studies the material, or equipment, as it were, that languages provide for communication.

In addition to *expression meaning*, there is *utterance meaning*. It comes about when a sentence with its meaning is actually used in a concrete context. First of all, *utterance meaning* involves *reference*. In addition to, and in connection with, reference another central notion comes into play, the notion of *truth*. The question of *truth* primarily concerns declarative sentences. Only such sentences, when uttered, are true or false. But it matters also for interrogative and other types of sentences. For example, if John asked Mary *Do you need my bicycle?*, the use of the question form would convey that he wants to know from his addressee whether, from her perspective, *I need your bicycle* is true or false.

In the literature of semantics, the notion of utterance meaning is not used in a uniform way. The technical term for this is *context of utterance*. Roughly speaking, the context of utterance is the sum of circumstances that bear on reference and truth. The most important ones are the following aspects: (a) the speaker (or producer) of the utterance, (b) the addressees (or recipients) of the utterance, (c) the time at which the utterance is produced and/or received, (d) the place where the utterance is produced and/or received, and e) the facts given when the utterance is produced and/or received.

In certain cases, such as communication by mail, the time, place, and facts may differ for the production of an utterance and its reception. Given this background utterance, meaning can be defined as the meaning that results from using an expression in a given communicative utterance. Utterance meaning derives from expression meaning on the basis of the particulars provided by the communicative utterance. The only aspects of the communicative utterance that matter are those that immediately bear on reference and truth of the expression.

Making Inferences

When someone produces an utterance, the addressees usually make various kinds of inferences. Although these inferences are somehow triggered in the addressee's mind, it is important to separate what is actually being said from what is inferred. Some authors prefer not to draw this distinction and adopt a very wide notion of utterance meaning. The investigation of such inferences, their role in communication and how they are related to the meaning, of what is actually said, is an important part of the linguistic discipline called *pragmatics*, the scientific study of, roughly speaking, the rules that govern language use. Within pragmatics, Paul Grice's theory of conversational implicatures deals with inferences of this kind.

Utterance meaning is also of concern for semantics: It has to explain how reference and truth depend on the context of utterance. For example, a semantic theory of tense would have to describe and explain which relations to the time of utterance present, past, and future tense forms can indicate. A further important subject, gradually gaining

importance, is the analysis of the systematic meaning shifts that expression meanings may undergo.

Communicative Meaning

Neither the level of expression meaning nor that of utterance meaning is the primary level on which we interpret verbal utterances. In an actual exchange, the main concern inevitably is this: What does the speaker intend with the utterance; in particular, what does the speaker want from me? Conversely, when we take on the speaking part, we choose our words in pursuit of a certain communicative intention. Verbal exchanges are a form of social interaction. They form an important part of our social lives. As such, they will always be interpreted as part of the whole social exchange and relationship entertained with the speaker.

A theory that addresses this level of interpretation is speech act theory, introduced in the 1950s by the philosopher John L. Austin and developed further by others, in particular John R. Searle. The central idea of speech act theory is that whenever we make an utterance in a verbal exchange we act on several levels. One level is what Austin calls the *locutionary act*. A locutionary act is the act of using a certain expression (usually a sentence) with a certain meaning in the given context of utterance. In doing so, we also perform an illocutionary act on the level on which the utterance constitutes a certain type of *speech act*: a statement, a question, a request, a promise, a refusal, a confirmation, a warning, and so forth.

The speech act level will be referred to as *communicative meaning*. Unlike expression meaning and utterance meaning, communicative meaning lies outside the range of semantics. Rather, it is of central concern for pragmatics.

Therefore, a more precise and comprehensive definition of semantics would be the following. Semantics is the study of the meanings of linguistic expressions, either simple or complex, taken in isolation. It further accounts for the way utterance meaning—that is, the meaning of an expression used in a concrete context of utterance—is related to expression meaning.

To summarize the levels of meaning, the following definitions can be surmised: (a) Expression meaning is the meaning of a simple or complex

expression taken in isolation; (b) utterance meaning is the meaning of an expression when used in a given context of utterance with fixed reference and truth value (for declarative sentences); and (c) communicative meaning is the meaning of an utterance as a communicative act in a given social setting. Hence, three levels of meaning are identified and defined.

Sentence Meaning and Compositionality

Although we usually understand sentences without any conscious effort, their meanings must be derived from our stored linguistic knowledge. This process is technically called *composition*. Complex expressions whose meanings are not stored in the lexicon are therefore said to have *compositional meaning*.

The meaning of complex expressions is determined by semantic composition. This mechanism draws on three sources: (a) the lexical meanings of the basic expressions; (b) the grammatical forms of the basic expressions; and (c) the syntactic structure of the complex expression. Thus, the principle of compositionality states that the meaning of a complex expression is determined by the lexical meanings of its components, their grammatical meanings, and the syntactic structure of the whole. The principle implies that the meanings of complex expressions are fully determined by the three sources mentioned—that is, by the linguistic input alone. Thus, in particular, the process does not draw on extralinguistic context knowledge. The principle as it stands can therefore be taken as an indirect definition of expression meaning, since expression meaning is that level of meaning that can be obtained by the process of composition—that is, on the basis of lexical meanings, interpretation rules for grammatical forms, and semantic combination rules.

It should be noted that the principle does not hold for the level of utterance meaning. Utterance meaning can only be determined by bringing in nonlinguistic knowledge about the given context of utterance.

Therefore, the principle of compositionality yields a convenient division of semantics into the following subdisciplines: *lexical semantics*, which is the investigation of expression meanings stored in the mental lexicon; *compositional word*

semantics, which is the investigation of the meanings of words that are formed by the rules of word formation; *semantics of grammatical forms*, which is the investigation of the meaning contribution of grammatical forms that can be freely chosen, often understood as including the semantic analysis of function words such as articles, prepositions and conjunctions; and *sentence semantics*, which is the investigation of the rules that determine how the meanings of the components of a complex expression interact and combine. Often semantics is subdivided into two subdisciplines only: Lexical semantics is then understood as also comprising compositional word semantics, and the semantics of grammatical forms is subsumed under sentence semantics.

Meanings as Concepts

The meaning is a mental description. For mental descriptions in general the term *concept* can be used. A concept for a kind, or category, of entities is information in the mind that allows us to discriminate entities of that kind from entities of other kinds. However, a concept should not be equated with a visual image. Many categories we have words for such as *mistake*, *thought*, *noise*, *structure*, and *mood* are not categories of visible things. But even for categories of visible things such as *dogs*, the mental description is by no means exhausted by a specification of their visual appearance. The *dog* concept, for instance, also specifies the behavior of dogs and how dogs may matter for us as pets, guide dogs, and so on.

The ideas and statements about the meanings of words and sentences can be presented as follows. The meaning of a word, more precisely, a content word (noun, verb, and adjective), is a concept that provides a mental description of a certain kind of entity. The meaning of a sentence is a concept that provides a mental description of a certain kind of situation.

Descriptive Meaning of Sentences: Propositions

The descriptive meaning of a sentence, its proposition, is a concept that provides a mental description of the kind of situations it potentially refers to. It is not only content words that shape the descriptive meaning of

the sentence. Functional elements, such as pronouns and articles or tense, a grammatical form, contribute to the proposition as well. Thus, the descriptive meaning of a word or a grammatical form is its contribution to descriptive sentence meaning.

Denotations and Truth Conditions: Semiotic Triangle

The denotation of a content word is the category, or set, of all its potential referents. The denotation of a word is more than the set of all existing entities of that kind. It includes real referents as well as fictitious ones, usual exemplars and unusual ones, maybe even exemplars we cannot imagine because they are yet to be invented.

The relationship between a word, its meaning, and its denotation is often depicted in the semiotic triangle, a convenient schema that is often used by linguists.

There is no established term for what would be the denotation of a sentence. By analogy with the denotation of a content word it would be the set, or category, of all situations to which the sentence can potentially refer; that is, the category of all situations in which the sentence is true. There is, however, another notion that is quite common and directly related to the would-be denotation of a sentence: Its so-called truth conditions. The truth conditions of a sentence are the conditions under which it is true. If we know the truth conditions of a sentence, then we know which situations the sentence can refer to; that is, the denotation of the sentence. Conversely, if we know which situations a sentence can refer to, we know its truth conditions. Thus, the notion of truth conditions can be considered the practical equivalent of the notion of the denotation of a sentence. The descriptive meaning of the sentence is its proposition, and the proposition determines the truth conditions of the sentence.

It should be also noted that the grammatical type of a sentence contributes its meaning, and this contribution is nondescriptive. For instance, the two propositions can be compared: (a) *The rain has ruined our trip*; (b) *Has the rain ruined our trip?* In this set of sentences, the question describes exactly the same sort of situation. Hence, it is considered to have the same proposition as the declarative sentence. The difference in

meaning is due to the grammatical forms of the sentences, or more precisely, to differences in grammatical sentence type. The semantic contribution of the grammatical sentence type is not part of the proposition. For declarative sentences, it consists in presenting the situation expressed as actually pertaining. This sentence type is therefore used for making assertions and communicating information. The interrogative sentence type, by contrast, leaves open whether or not the situation pertains. It is therefore the standard option to be chosen for asking questions. Thus, the meaning contribution of grammatical sentence type is non-descriptive meaning.

Linguistic Approach to Semantics

The key element of the modern linguistic approach to semantics is that there is no escape from language. The goal of semantics is to study relations within language such as paraphrase or synonymy, paraphrase or antonymy, and so forth. Another task of semantics is seeing language study as the explication of the linguistic competence of the native speaker of a language, or in other words the provision of rules and structures which specify the mental apparatus a person must possess if he is to know a given language. For the linguist, the main difficulty lies in drawing a boundary not simply between sense and nonsense but between the kind of nonsense that arises from contradicting what we know about language and meaning, and the kind of nonsense that comes from contradicting what we know about the real world.

Mariam Orkodashvili

See also Artificial Intelligence; Information Theory; Information, Qualitative; Information, Quantitative; Linguistic Frameworks; Linguistics, Historical; Semantics (Scientific and Empirical); Syntax (Introduction to Theory); Syntax (Scientific and Empirical)

Further Readings

- Aarts, B. (2007). *Syntactic gradience. The nature of grammatical indeterminacy*. Oxford University Press, 2007.
- Aarts, B., Denison, D., Keizer, E., & Popova, G. (Eds.). (2004). *Fuzzy grammar*. Oxford University Press.

- Biber, D., Johansson, S., Leech, G., Conrad, S., & Finegan, E. (1999). *Longman grammar of spoken and written English*. Pearson Education Ltd.
- Biber, D., Reppen, R., & Conrad, S. (2004). *Corpus linguistics. Investigating language structure and use*. Cambridge University Press.
- Blakemore, D. (2004). *Relevance and linguistic meaning. The semantics and pragmatics of discourse markers*. Cambridge University Press.
- Bloomfield, L. (1984/1933). *Language*. University of Chicago Press.
- Boeckx, C. (2006). *Linguistic minimalism. Origins, concepts, methods, and aims*. Oxford University Press.
- Chomsky, N. (1957). *Syntactic structures*. Mouton.
- Chomsky, N. (1981). *Lectures on government and binding*. Foris.
- Chomsky, N. (2002). *New horizons in the study of language and mind*. Cambridge University Press.
- Chomsky, N. (2007). *On language*. The New Press.
- Cook, G., & Seidlhofer, B. (Eds.). *Principle and practice in applied linguistics*. Oxford University Press.
- Crystal, D. (Ed.). (1995). *The Cambridge encyclopedia of the English language*. Cambridge University Press.
- Crystal, D. (Ed.). (1997). *The Cambridge encyclopedia of language*. Cambridge University Press.
- Dancygier, B. (2006). *Conditionals and prediction. Time, knowledge and causation in conditional constructions*. Cambridge University Press.
- Dancygier, B., & Sweetser, E. (2009). *Mental spaces in grammar. Conditional constructions*. Cambridge University Press.
- de Saussure, F. (2011). *Course in general linguistics* (P. Meisel & H. Saussy, Eds. & W. Baskin, Trans.). Columbia University Press.
- de Saussure, F. (1931). *Cours de linguistique générale*. Payot.
- Haegeman, L. (2005). *The syntax of negation*. Cambridge University Press.
- Halliday, M. A. K., & Hasan, R. (1976). *Cohesion in English*. Routledge.
- Hart, C., & Lukes, D. (Eds.). (2007). *Cognitive linguistics in critical discourse analysis. Application and theory*. Cambridge Scholars Publishing.
- Von Heusinger, K. (2002). Specificity and definiteness in sentence and discourse structure. *Journal of Semantics*, 19.
- Hudson, R. (1999). *Sociolinguistics*. Cambridge University Press.
- Jackendoff, R. (2009). *Language, consciousness, culture. Essays on mental structure*. MIT Press.
- Jespersen, O. (1964). *Language: Its nature, development, and origin*. Norton.
- Johnson, K. (2008). *Quantitative methods in linguistics*. Blackwell.
- Kager, R. (1999). *Optimality theory*. Cambridge University Press.
- Karttunen, L. (1976). Discourse Referents. In J. McCawley (Ed.), *Syntax and semantics 7: Notes from the linguistic underground*. Academic Press.
- Kempson, R. M. (1977). *Semantic theory*. Cambridge University Press.
- Labov, W. (2001). *Principles of linguistic change*. Blackwell.
- Lakoff, G. (1987). *Women, fire, and dangerous things. What categories reveal about the mind*. University of Chicago Press.
- Leech, G. (1983). *Semantics. The study of meaning*. Penguin Books.
- Löbner, S. (2002). *Understanding semantics*. Oxford University Press.
- Lyons, J. (1972). *Introduction to theoretical linguistics*. Cambridge University Press.
- Lyons, J. (1995). *Linguistic semantics. An introduction*. Cambridge University Press.
- Meillet, A. (1948). *Linguistique historique et linguistique générale [Historic and general linguistics]*. Champion.
- Niyogi, P. (2006). *The computational nature of language learning and evolution*. MIT Press.
- Parikh, P. (2010). *Language and equilibrium*. MIT Press.
- Quirk, R., Greenbaum, S., Leech, G., & Svartvik, J. (Eds.). (2000). *A comprehensive grammar of the English language*. Pearson Education Limited.
- Sapir, E. (1921). *Language*. Harcourt, Brace and World.
- Whorf, B. L. (1956). The relation of habitual thought and behavior to language. In J. B. Carroll (Ed.), *Language, thought, and reality: Selected writings of Benjamin Lee Whorf* (pp. 134–159). MIT Press.
- Widdowson, H. G. (1996). *Linguistics*. Oxford University Press.
- Woods, A. (1986). *Statistics in language studies. Cambridge textbooks in linguistics*. Cambridge University Press.

SEMANTICS (SCIENTIFIC AND EMPIRICAL)

Semantics is the aspect of the discipline of linguistics that is concerned with meaning. This entry discusses some of the more technical aspects of

semantics. For a general introduction, please see the entry Semantics (Introduction to Theory) in this encyclopedia.

The sections that follow present the fundamentals of key topics in semantics, including lexemes, homonyms, polysemy, and semantic ambiguity. Next, expressive meaning and the importance of cognitive psychology for semantics are discussed, and the entry concludes with some observations on the relationship between meaning and cognition.

Lexemes and Lexicon of a Language

Lexemes are focal elements of modern semantic analysis, both scientific and empirical. Expressions with a lexical meaning are referred to as *lexemes* or *lexical items*. Lexemes constitute the lexicon of the language, a huge complex structure in the minds of the language users. Lexical meanings we have in our mental lexicons are not just paraphrases. They are concepts.

Lexemes differ in their grammatical behavior and are accordingly assigned to different grammatical categories such as verbs, nouns, and adjectives. Occasionally the term *word class* is used instead of grammatical category. The grammatical category determines the way in which a lexeme can be used within a sentence. For example, a noun can be combined with an adjective, and the whole with an article to form a noun phrase. The noun phrase in turn can be combined as object with a verb to form a verb phrase, and so forth.

Thus, a lexeme is a linguistic item defined by the following specifications, which make up what is called the *lexical entry* for this item: (a) its sound form and its spelling (for languages with a written standard); (b) the grammatical category of the lexeme (noun, intransitive verb, adjective, etc.); (c) its inherent grammatical properties (for some languages, e.g., gender); (d) the set of grammatical forms it may take, in particular irregular forms; and (e) its lexical meaning. These specifications apply both to simple and composite lexemes. Composite lexemes too, such as *throw in the towel* or *red light*, have a fixed sound form, spelling, and lexical meaning. They belong to a grammatical category (intransitive verb for *throw in the towel* and noun for *red light*), they have inherent grammatical properties and the usual range of grammatical forms. For example, the grammatical

forms of *throw in the towel* are those obtained by inserting the grammatical forms of the verb *throw*: *throws in the towel*, *threw in the towel*, and so on.

In principle, each of the specifications of a lexeme is essential: If two linguistic items differ in one of these points, they are considered different lexemes. There are, however, exceptions. Some lexemes have orthographic variants, for example, *rhyme/rime*, others may have different sound forms, for example, *laboratory* may be stressed on either the first or the second syllable. The American and the British variants of English differ in spelling and sound form for many lexemes. As long as all the other properties of two orthographic or phonetic variants of a lexeme are identical, in particular their meanings, they will not be considered different lexemes but lexemes with a certain limited degree of variation. That does not mean that minor differences do not count. They count in any event if they are connected with a difference in meaning.

Homonymy, Polysemy, and Vagueness

Homonymy

Words with the same historical origin are called *cognates*. *Light* is a cognate of a different German word, the adjective *leicht* (light, easy). Due to their different origins, these two units are considered two different lexemes by most linguists. In general, different meanings are assigned to different lexemes if they have different historical sources. The idea is that as long as their meanings remain distinct, different words do not develop into one, even if their sound forms and/or spellings happen to coincide for independent reasons. The two adjectives *light* and *leicht* are an instance of what is called *total homonymy*: two lexemes that share all distinctive properties (grammatical category and grammatical properties, the set of grammatical forms, sound form, and spelling), yet have unrelated different meanings. One would talk of partial homonymy if two lexemes with different unrelated meanings coincide in some but not all of their grammatical forms, for instance, the verbs *lie* (*lay*, *lain*) and *lie* (*lied*, *lied*). Partial homonyms can give rise to ambiguity in some contexts (do not lie in bed) but can be distinguished in others (*he lay/lied in bed*).

Homonymy can be related either to the sound forms of the lexemes or to their spellings: Homonymy with respect to the written form is *homography*; if two lexemes with unrelated meanings have the same sound form, they constitute a case of *homophony*.

Polysemy

A lexeme constitutes a case of polysemy if it has two or more interrelated meanings, or better: meaning variants. The phenomenon of polysemy is independent of homonymy: of two homonyms, each can be polysemous. It results from a natural economic tendency of language. Rather than inventing new expressions for new objects, activities, experiences, and so forth to be denoted, language communities usually opt for applying existing terms to new objects, terms hitherto used for similar things. Scientific terminology is one source contributing to polysemy on a greater scale. Some scientific terms are newly coined, but most of them will be derived from ordinary language use. Among the terms introduced in semantics, *lexeme*, *homonymy*, and *polysemy* are original scientific terms, while others, such as *meaning*, *reference*, and *composition*, are ordinary expressions for which an additional technical meaning variant was introduced. In principle, polysemy is a matter of single lexemes in single languages. Polysemy plays a major role in the historical development of word meanings because lexemes continually shift their meanings and develop new variants.

Semantic Ambiguity or Vagueness

Many words have more than one meaning, and even complete sentences may allow for several readings. The term for this phenomenon is *ambiguity* (or *vagueness*): An expression or an utterance is ambiguous if it can be interpreted in more than one way. The notion of ambiguity can be applied to all levels of meanings: expression meaning, utterance meaning, and communicative meaning. Widespread vagueness in the lexicon should be considered another economic trait of language. Vagueness may occur in combination with polysemy. For example, the meaning variants of *light* are a matter of different underlying scales of weight, difficulty, and so on. Each meaning of scale is vague.

Synonymy

Two lexemes are synonymous if they have the same meaning. Synonymy in the strict sense includes all meaning variants for two polysemous lexemes and it includes all meaning parts; namely, descriptive, social, and expressive meaning.

Meaning and Subjectivity: Expressive Meaning

Anything we say will not only be understood through its denotation or descriptive meaning but also be taken as the expression of a personal emotion, opinion, or attitude. However, a sentence can be spoken in a neutral manner or in a way that exhibits different emotions, for example, desperation or concern but also delight, amusement, or relief. It could be pronounced in a mean or a kind manner, or hastily, or particularly slowly. This general aspect of language use, its expressive function, is not what is meant with the notion of expressive meaning. On a par with descriptive and social meaning, expressive meaning is part of the lexical meaning of certain expressions, a semantic quality of words and phrases independent of the context of utterance and of the way they are being spoken.

The expressions with expressive meanings are called, unsurprisingly, *expressives*. Again, there are two kinds of expressive, those with exclusively expressive meaning and others with both descriptive and expressive meaning. The most typical instances of expressives with exclusively expressive meaning are words and phrases used for directly expressing an emotion, feeling, or sensation, such as *ouch*, *wow*, and *oh*. Such interjections are language-specific. Languages may differ in how many such expressions they have, and an expressive may have different meanings in different languages.

At least some feelings, sensations, attitudes, and evaluations can thus be expressed in two ways: subjectively and directly by means of expressives and objectively by forming sentences with the respective descriptive meaning. For instance, *Ouch!* and *That hurts!* While expressing similar emotions, they are structurally different, the former being the expressive interjection and the latter being an ordinary short exclamation describing the feeling.

The rules governing the use of expressives are simple. Since all expressives serve to express personal feelings, attitudes, or sensations which are perceptible only to the holder, their correct use is just a matter of personal judgment. Interjections and exclamations can be used as complete utterances. Other expressives such as *hopefully*, *(un)fortunately*, or *thank God* can be inserted into a sentence in order to add a personal attitude to the situation expressed.

Meaning and Cognition

During the past three decades, the development of a new branch of science, cognitive psychology, or more generally cognitive science, has given important fresh impetus to linguistics in general and semantics in particular. Cognitive science is concerned with how the human mind works, how it receives information from the environment via the senses and processes this information, recognizing what is perceived, comparing it to former data, classifying it, and storing it in the memory. It tries to account for the complex ways in which the vast amount of information in our minds is structured, and how we can operate with it when we think and reason. Language plays a central role in these theories. On one hand, speech perception and production and the underlying mental structures are major objects of investigation. On the other hand, the way in which we use language to express what we have in mind can tell much about how the human mind is organized.

The importance of cognitive psychology for semantics lies in its emphasis on the exploration of the concepts and categories we use. While semantics in the wider tradition of structuralism aims at a description of meaning relations, the cognitive approach focuses on meanings themselves. It tries to provide positive descriptions of word meanings and thereby to account for why and how words denote what they denote.

The mental act of classifying things and forming categories which underlies cognition in all its aspects is called *categorization*. Traditionally, categorization was held to be a matter of necessary and sufficient conditions (e.g., a bird belongs to the category of penguins if and only if it fulfills the

particular set of conditions that define this category).

This view was challenged by the proponents of so-called *prototype theory*. These cognitivists argued that categories are essentially defined by prototypes—in other words, those examples of a category that come to mind first and represent the most typical cases. Prototype theory led to a radically different view of categorization implying, for example, fuzzy category boundaries and gradual membership. Another significant novelty was the investigation of the hierarchical organization of our category system, with the result of establishing what came to be called the *basic level of categorization*. This is a medium level of generality that is preferred in thinking and communicating. However, many of the central claims of prototype theory, in particular, fuzzy category boundaries and gradual membership, are problematic.

Categorization is only possible if the respective categories are in some way available in the cognitive system; that is, in our mind. In the terminology of cognitive science, categorization requires mental representations of the categories. Categories are represented by concepts for their exemplars. When we encounter an object, our cognitive apparatus will produce a preliminary description of it, which consists of what we perceive of the object: its size, shape, color, smell, and other such qualities. The description will be compared with the concepts we have in memory, and if the description happens to match with the concept, the object will be categorized as such, under the relevant category.

Another important issue is the relationship between meaning and cognition: the distinction between the meaning of a word and the total knowledge we mentally connect to the word. More specifically, this distinction between semantic knowledge and cultural knowledge is important for semantics.

Mariam Orkodashvili

See also Artificial Intelligence; Information, Qualitative; Information, Quantitative; Linguistic Frameworks; Linguistics, Historical; Semantics (Introduction to Theory); Statistics; Syntax, (Introduction to Theory); Syntax (Scientific and Empirical)

Further Readings

- Aarts, B. (2007). *Syntactic gradience. The nature of grammatical indeterminacy*. Oxford University Press.
- Aarts, B., Denison, D., Keizer, E., & Popova, G. (Eds.). (2004). *Fuzzy grammar*. Oxford University Press.
- Biber, D., Johansson, S., Leech, G., Conrad, S., & Finegan, E. (1999). *Longman grammar of spoken and written English*. Pearson Education Ltd.
- Biber, D., Reppen, R., & Conrad, S. (2004). *Corpus linguistics. Investigating language structure and use*. Cambridge University Press.
- Blakemore, D. (2004). *Relevance and linguistic meaning. The semantics and pragmatics of discourse markers*. Cambridge University Press.
- Bloomfield, L. (1984/1933). *Language*. University of Chicago Press.
- Boeckx, C. (2006). *Linguistic minimalism. Origins, concepts, methods, and aims*. Oxford University Press.
- Chomsky, N. (1957). *Syntactic structures*. Mouton.
- Chomsky, N. (1981). *Lectures on government and binding*. Foris.
- Chomsky, N. (2002). *New horizons in the study of language and mind*. Cambridge University Press.
- Chomsky, N. (2007). *On language*. The New Press.
- Cook, G., & Seidlhofer, B. (Eds.). (1995). *Principle and practice in applied linguistics*. Oxford University Press.
- Crystal, D. (Ed.). (1995). *The Cambridge encyclopedia of the English language*. Cambridge University Press.
- Crystal, D. (Ed.). (1997). *The Cambridge encyclopedia of language*. Cambridge University Press.
- Dancygier, B. (2006). *Conditionals and prediction. Time, knowledge and causation in conditional constructions*. Cambridge University Press 1995.
- Dancygier, B., & Sweetser, E. (2009). *Mental spaces in grammar. Conditional constructions*. Cambridge University Press.
- de Saussure, F. (2011). *Course in general linguistics* (P. Meisel & H. Saussy, Eds. & W. Baskin, Trans.). Columbia University Press.
- de Saussure, F. (1931). *Cours de linguistique générale*. Payot.
- Haegeman, L. (2005). *The syntax of negation*. Cambridge University Press.
- Halliday, M. A. K., & Hasan, R. (1976). *Cohesion in English*. Routledge.
- Hart, C., & Lukes, D. (Eds.). (2007). *Cognitive linguistics in critical discourse analysis. Application and theory*. Cambridge Scholars Publishing.
- Hudson, R. (1999). *Sociolinguistics*. Cambridge University Press.
- Jackendoff, R. (2009). *Language, consciousness, culture. Essays on mental structure*. MIT Press.
- Jespersen, O. (1964). *Language: Its nature, development, and origin*. Norton.
- Johnson, K. (2008). *Quantitative methods in linguistics*. Blackwell.
- Kager, R. (1999). *Optimality theory*. Cambridge University Press.
- Karttunen, L. (1976). Discourse referents. In J. McCawley (Ed.), *Syntax and semantics 7: Notes from the linguistic underground*. Academic Press.
- Kempson, R. M. (1977). *Semantic theory*. Cambridge University Press.
- Labov, W. (2001). *Principles of linguistic change*. Blackwell.
- Lakoff, G. (1987). *Women, fire, and dangerous things. What categories reveal about the mind*. University of Chicago Press.
- Leech, G. (1983). *Semantics. The study of meaning*. Penguin Books.
- Löbner, S. (2002). *Understanding semantics*. Oxford University Press.
- Lyons, J. (1972). *Introduction to theoretical linguistics*. Cambridge University Press.
- Lyons, J. (1995). *Linguistic semantics. An introduction*. Cambridge University Press.
- Meillet, A. (1948). *Linguistique historique et linguistique générale [Historic and general linguistics]*. Champion.
- Niyogi, P. (2006). *The computational nature of language learning and evolution*. MIT Press.
- Parikh, P. (2010). *Language and equilibrium*. MIT Press.
- Quirk, R., Greenbaum, S., Leech, G., & Svartvik, J. (Eds.) (2000). *A Comprehensive grammar of the English language*. Pearson Education Limited.
- Sapir, E. (1921). *Language*. Harcourt, Brace and World.
- Von Stechow, K. (2002). Specificity and definiteness in sentence and discourse structure. *Journal of Semantics*, 19.
- Whorf, B. L. (1956). The relation of habitual thought and behavior to language. In J. B. Carroll (Ed.), *Language, thought, and reality: Selected writings of Benjamin Lee Whorf* (pp. 134–159). MIT Press.
- Widdowson, H. G. (1996). *Linguistics*. Oxford University Press.
- Woods, A. (1986). *Statistics in language studies. Cambridge textbooks in linguistics*. Cambridge University Press.

SET THEORY

Set theory is the study of membership in collections of elements or *sets*. Beyond this, set theory can be understood as a symbolic framework with wide applications outside mathematics, yet which

has been applied by mathematicians in two ways: First, set theory can be seen as an approach to advance knowledge of particular mathematical problems, notably as the study of the formal properties of infinity. Second, set theory is seen as a language that allows all of mathematics to be subsumed under one unifying theory.

This entry is divided into three sections. The first explains the core concepts of set theory, including notation and the concept of *cardinality*, or the size of sets. The second section outlines the invention of modern set theory by Georg Cantor (1845–1918) and his counterintuitive discoveries, notably that there are different *sizes* of infinity. The third section describes the efforts by scholars to use set theory as a foundation of mathematics. Finally, the entry summarizes the philosophical division of mathematics raised by set theory: whether mathematics can be understood as the study of something *real* or whether it is a symbol system that is invented or constructed by the human mind.

Despite these debates, and the paradoxes around mathematics, set theory continues to remain as the dominant theory used to unify disparate areas of mathematical inquiry. Moreover, the startling finding by Cantor that infinity consists of different *sizes* is noncontroversial. To begin understanding set theory, it is important to understand some of its basic concepts.

Set Theory as a Formal Language

Set theory is a formal language to define membership relations, as opposed to a natural language.

Formal languages are useful because they specify collections of objects precisely, that is, without ambiguity. Without precise definitions, it is more difficult to communicate abstract relationships and ideas. Applications of set theory are useful for understanding many areas of mathematics and computer science, where precise terms are needed to define objects of inquiry.

As noted in the introduction, a *set* is a collection of distinct *elements*. Elements of a set can be anything, including other sets or nothing at all. Sets with an infinite number of elements are called *infinite sets*, and sets with a finite number of elements are called *finite sets*. Sets are typically represented as capital letters, with a few exceptions. In particular, sets with no elements, or *empty sets*, are denoted by the $\emptyset$ symbol. Other common set theory terms are listed in Table 1.

Elements are listed and separated by commas within curly brackets “{ }” that denote the upper and lower bounds of the set. Ellipses indicate a continuous of a sequence within a set. In addition to listing elements of a set, the *set builder* notation, typically notated using a vertical bar “|” can be used to describe the criterion of a set.

Set theory also uses a stylized form of the Greek *epsilon*, $\in$, to express set membership (*epsilon* is the first letter of the Greek word “is”). Thus, “ $x \in A$ ” means that “ x is an element of A .” If *epsilon* is reversed ($\ni$), then the expression is “ $A \ni x$,” which means that “set A has x as an element.” Conversely, “ $x \notin A$ ” means that “ x is not an element of set A ,” and “ $A \not\ni x$ ” means that “set A does not have x as an element.”

Table 1 Common Set Theory Terms

Symbol	Meaning	Example	Translation
{,...}	set; collection of elements	$A = \{1, 2, 3\}$ $B = \{3, 4, 5\}$	“The set A of elements 1, 2, and 3.” “The set B of elements 3, 4, and 5.”
$\in$	“belongs to”	$1 \in A$	“The element 1 belongs to set A .”
$\notin$	“does not belong to”	$4 \notin A$	“The element 4 does not belong to set A .”
$\ni$	“contains”	$A \ni 1$	“Set A contains element 1.”
$\not\ni$	“does not contain”	$A \not\ni 4$	“Set A does not contain element 4.”
	“such that”	$\{x \mid x > 0\}$	“All x such that x is greater than zero.”

Note: Terms used are to identify membership and their negation. The symbol “|” is also known as *set-builder notation*.

Source: Compiled by Nathan E. Fosse.

Some additional examples of the foregoing nomenclature are shown here. For example, $A = \{1, 2, 3, 4, 5\}$ can be used to mean “the set of the set all integers from one to three.” Sets can have all elements listed or can be described as a criterion for membership. For example, the following set A consists of five elements, wherein $B = \{\text{Liam, Noah, Ethan, Mason, Lucas}\}$. This can also be described as, $B = \{x \mid x \text{ is one of the five most popular boy names in the year 2015}\}$. Alternatively, this set is typically expressed as, “the set B of all x , such that x is one of the five most popular boy names in the year 2015.” As the previous examples show, set theory is a way of defining collections of elements and their relative membership. In this way, sets are useful for clarifying natural language (or everyday language) in more precise terms, that is, by translating natural language into a formal language. Set theory is also a way of linking logical statements and mathematical statements under the same system of language.

A core concept in the study of sets is the determination of their size or the number of elements that sets contain. The absolute number of elements in a set is called its *cardinality*. The relative size of sets is called their *ordinality*. Put differently, the cardinality of a set is the number of elements it contains; in contrast, the ordinality of a set is whether or not it contains more, fewer, or the same number of elements as another set. Thus, two sets are considered the same size or of the same cardinality if they contain the same number of elements. For example, set $A = \text{set } B$ because they exhibit the same size or cardinality; in finite sets, this can be simply a means of counting the number of elements to determine their cardinality. Another way to compare sizes of sets is by their *ordinality*, that is, by whether one is relatively larger than another. Thus, while *cardinality* refers to absolute size of sets, *ordinality* is the *relative* size of sets. The size of sets is the number of elements in them. For finite sets, ordinality is the same as cardinality: That is, the number of elements that are in a set define whether it is larger, smaller, or the same size as another set.

A central insight in set theory is that we can determine whether two sets are of equal size if we can *pair up* each element of one set to another; in set-theoretic terms, two sets are of equal size (the same *cardinality*) if their elements have a

Table 2 One-to-One Correspondence of Two Finite Sets, A and B

A	1	2	3	4	5
	$\updownarrow$	$\updownarrow$	$\updownarrow$	$\updownarrow$	$\updownarrow$
B	Liam	Noah	Ethan	Mason	Lucas

Note: “One-to-one correspondence” means sets compared are the same size (cardinality).

Source: Compiled by Nathan E. Fosse.

one-to-one correspondence. The example is clear with finite sets, although the crucial insight in the study of infinity is that this *matching* technique can be applied to determine the size of infinite sets. The intuition of one-to-one correspondence is clearer when comparing sets A and B as above.

Set membership can be explained in terms of *subsets*. *Subsets* are those sets where every element is in another set. Following this definition, a subset can be equal to a set that contains the same items. Moreover, because every element of an empty set is an element in any set of finite numbers S , every set of finite numbers has as its subset the empty set $\emptyset$.

Two other types of subsets are *proper subsets* and *power subsets*. First, proper subsets are more specific. In set theory, a *proper subset*, for example, is one where K is a subset of A but where the number of elements in K is not equal to A . For example, the set of all Kansans is a proper subset of all Americans. Kansans are a subset of Americans because every person in Kansas is also an American, but the number of Kansans does not equal the number of Americans (in all 50 states). As mentioned previously, elements of sets can be anything, including other sets. Each set not only contains subsets but also a *powerset* that is the set of all subsets in S . Powersets are indicated by the Greek symbol delta (lowercase), δ ; for example, for set S , $\delta(S)$ is the set of all powersets for S . Determining the cardinality of the powerset for any set S with n elements is straightforward: $\delta S = 2^n$. For example, given the example above, set A contains five elements. Accordingly, the powerset of A is 32; that is, there are 32 subsets of set A . Formally, $\delta A = 2^5 = 32$.

Another fundamental set of logical operations between sets concerns their degree of membership

with each other. Three of the most common operations between sets are described below. Below are several of the essential operations in set theory. First, the *union* of sets A and B ($A \cup B$) includes membership to all distinct elements of A and B . Second, the *intersection* of two sets A and B ($A \cap B$) restricts membership to only those elements belonging to A and B jointly, but not to any other elements. Third, the *complement* of set A (also known as the *absolute complement*) refers to anything not belonging to that set A (denoted as A^c). Extensions of the union, complement, and interaction of sets are used to define set membership.

The present treatment is by no means exhaustive; instead, basic concepts were introduced to clarify the language necessary to describe set membership and the relative sizes of sets. A comprehensive source is Paul R. Halmos's 1974 book *Naïve Set Theory*. For a more accessible introduction to set theory and related concepts, see Jan Gullberg's *Mathematics: From the Birth of Numbers* (1997).

The next section describes these applications for infinite sets. In fact, set theory was invented by the German mathematician Georg Cantor for the purpose of studying the relative size of infinity. Before Cantor, there was only a vague concept of infinity but no formal study of its properties. These insights, as will be shown, are counterintuitive and were controversial in their time. However, Cantor's application of set theory was the catalyst for the search for a unifying theory of foundation of mathematics based in set theory.

Set Theory as the Science of Infinity

Set theory is closely linked with the study of the infinite. That is, whereas Cantor believed that set theory is indeed the foundation of mathematics, he invented set theory to develop what he called an "arithmetic of the transfinite numbers" (quoted in Dauben, 1983, p. 112). Accordingly, the beginning of set theory is often marked by the groundbreaking 1874 publication by Cantor, the first published work proving the existence of different *sizes* of infinite number sets.

To understand set theory as the study of infinity, it is important to review different types of number sets and their notation. These types of numbers, their set-theoretic notation, description, and formal definition, are summarized in Table 3.

In particular, the following are typically used to reference number types: N , natural numbers; Z , integers; Q , rational numbers; I , irrational numbers; R , real numbers (notation can also rely on capital letters to describe these number sets: N , Z , Q , I , and R). Other number sets exist (such as complex or imaginary numbers), but these are described and referenced below because they most clearly demonstrate the counterintuitive insights from Cantor's discoveries.

Table 3 summarizes and describes the infinite sets using both natural language and its formal, set-theoretic definition. Note that all of these number sets are *infinite sets*; they continue in a never-ending sequence. Leaving out irrational numbers, let us compare the *ordinality*, or the relative size of these sets, based on the core concepts in the previous section (namely the concept of proper subset and of the union of sets). Taking an intuitive approach to infinite sets, it suggests sets where $N < Z < Q < R$. That is, N is a proper subset of Z , Z is a proper subset of Q , and Q is a proper subset of R . R consists of all possible numbers on the real number line. The informal, verbal description illustrates how these infinite sets intuitively relate to one another: all the following number sets— Z , Q , and R —build successively on the elements of N . To elaborate, Z is the union of natural numbers with positive whole integers and zero. Next, Q is the union of Z (the set of integers) and rational fractions; this is because every integer can be expressed as a fraction: $1/1 = 1$, $2/1 = 2$, $3/1$, and so on. Finally, because the set R is defined as the union of both rational and irrational numbers, it by definition includes Q and its proper subsets.

An intuitive approach to the proper subsets of the infinite set of natural numbers also suggests it contains proper subsets. For example, the number of positive even numbers, $E = \{2, 4, 6, 8, 10 \dots\}$ is, by definition, one half of all positive whole integers. Intuitively, there are fewer even numbers than natural numbers. In addition, one may intuitively conclude that there are fewer odd numbers, squared numbers, and primes; certainly, these are proper subsets if we restrict ourselves a finite set of natural numbers.

Cantor proved, however, that our intuition, based on the study of finite sets, cannot be true. The aforementioned concepts are intuitive when applying to finite sets, but when comparing sizes

Table 3 Common Infinite Number Sets

<i>Notation</i>	<i>Number Set</i>	<i>Informal Description</i>	<i>Formal Definition</i>
N	Natural numbers ¹	Whole positive numbers.	$N = \{1, 2, 3, \dots\}$
Z	Integers	Natural numbers plus their negative counterparts.	$Z = \{\dots -2, -1, 0, 1, 2, \dots\}$
Q	Rational numbers	Integers plus the rational fractions (those that can be expressed as a ratio of two integers).	$Q = \{ p/q : p \text{ and } q \text{ are members of } Z \text{ and } q \text{ is not zero} \}$
I	Irrational numbers	Fractions that <i>cannot</i> be expressed as a ratio of two integers; or, any real number that is not a rational number.	$I = \{ R - Q \}$
R	Real numbers	Rational numbers plus irrational numbers; the continuum all possible numbers on the number line.	$R = \{ I \cup Q \}$

Note: Natural numbers sometimes exclude zero. The essential point is that these numbers are *countable*. Zero is excluded here because of how natural numbers are used in set theory to count the number of elements in a set.

Source: Compiled by Nathan E. Fosse.

of infinite sets we must establish cardinality (absolute size) and ordinality (relative size) by pairing off elements of each set. Recall that two sets are the same size or have the same cardinality if they have a one-to-one correspondence. For infinite sets, we need not pair off *every* element, we only need to show that it is possible to establish a one-to-one correspondence.

By relying on the simple tool of one-to-one correspondences, Cantor proved there are two different levels or sizes of infinity. First, there are *countably infinite sets* that can be paired-off to the set of natural numbers N . These sets are of the same cardinality of N and are described as *countably infinite* because, at least in theory, one is able to count off each element of these infinite sets. The cardinality of these is defined by Cantor as *aleph-null* or $\aleph_0$ (*aleph* being the first letter of the Hebrew alphabet). The *null* subscript is to indicate that these infinite sets are the *smallest* size of infinity.

Second, Cantor found that other infinite sets are *uncountably infinite* because they cannot possibly be paired off with the set of real numbers. Uncountably infinite sets are vastly larger than the cardinality of countably infinite sets, $\aleph_0$. How much larger? Because we cannot establish the cardinality of uncountably infinite sets, we have no way of knowing how much larger, except that we

know that the difference is vast: One should think of countably infinite sets as minuscule compared to uncountably infinite sets. In fact, if we subtract from an uncountably infinite set the cardinality of countably finite sets, $\aleph_0$, *we are still left with an uncountably infinite set*.

Starting with countably infinite sets, the intuition is relatively simple. Beginning with the set of real numbers N , Cantor devised several techniques to *pair off* infinite sets with N . Contrary to intuition, Cantor proved that proper infinite subsets of N , such as the set of even numbers E , odd numbers O , squared numbers S , and prime numbers P , have a one-to-one correspondence with N . In general, there are many other countably infinite proper subsets that are the same size as N ; the point is that, because they have a one-to-one correspondence with N , we can establish that these have the same size or *cardinality*.

Moving beyond proper subsets of N , Cantor developed several ways of establishing a one-to-one correspondence between N and sets that are seemingly larger than N . In particular, by splitting the set of integers Z into odd and even numbers, Cantor showed that every even number and every odd number in Z can be paired with N , despite the fact that N is defined as a proper subset of Z . In the world of infinite sets, N is the same cardinality as Z , as Table 5 demonstrates.

Table 4 One-to-One Correspondences of Natural Numbers N and Proper Subsets

Natural numbers	N	1	2	3	4	...	n	...
Odd naturals	O	1	3	5	9	...	$2n - 1$	...
Even naturals	E	2	4	6	8	...	$2n$	...
Squared naturals	S	1	4	9	8	...	n^2	...
Prime naturals	P	3	5	7	11	...	n th prime	...

Note: "One-to-one correspondence" means sets compared are the same size (cardinality).

Source: Compiled by Nathan E. Fosse.

Table 5 One-to-One Correspondence of Infinite Sets N and Z

Natural numbers	N	1	2	3	4	...	$2n - 1$	$2n$	...
Integers	Z	0	1	-1	2	...	$-n + 1$	n	...

Note: "One-to-one correspondence" means sets compared are the same size (cardinality).

Source: Compiled by Nathan E. Fosse.

What about N and other seemingly larger sets, such as Q , the set of fractions? Here, Cantor again was able to establish a one-to-one correspondence, although this requires more complex diagrams; this is often called the *zigzag proof*. Cantor also showed that the set $N \times N = N$, in a complex diagram called the *zigzag proof*. (Illustrations of these are available in Yanovsky, pp. 74–75.)

In summary, *countably infinite sets* were those discovered by Cantor to have a one-to-one correspondence with the set of natural numbers N . By showing that these sets are of the same size or cardinality, Cantor invented a new language to examine the properties of the infinite.

If Cantor had stopped at the countably infinite sets, he would have found it impossible to prove different sizes of infinity. In fact, Cantor tried but

could not find a one-to-one correspondence between N and the set of all real numbers R (which Cantor also referred to R as the *continuum* or c). Instead of establishing the cardinality of R , Cantor proved that no such possible correspondence can possibly exist between N and R . That is, R is uncountably infinite and N is countably infinite; consequently, R is immensely larger than N . This discovery reveals that, numbers that are rarely identified, notably irrational and so-called *transcendent numbers* (i.e., numbers such as π that are both irrational and nonalgebraic) are far from rare: These numbers occupy most of the real number line.

Cantor proved that the set of real numbers R is uncountably infinite, using the famous *diagonalization proof*. This proof is illustrated in Table 6 and is a famous example of *proof by contradiction* (i.e., starting with a false premise and proving that such a premise cannot be true). The proof proceeds in several steps. First, we define the set of real numbers as the infinite set of real numbers in the interval $(0, 1)$. Each element in the set R , bound by 0 and 1, is itself an infinite sequence of decimals (the position of each decimal is indexed in the columns of Table 6). Second, we assume (incorrectly) that we can pair each element of R we have created with the set of natural numbers N . Thus, each row in Table 6 consists of a possible one-to-one correspondence of a real number and a natural number. Third, we show that, by changing the values of the diagonal numbers by one value (i.e., by changing out .4 to .5, and .9 to 0, and so on) we have now shown that a new number, D , is not in correspondence with the set of natural numbers. In fact, no matter what we do, we can *always* change the values of the diagonals by some increment, proving that the initial premise that $N = R$ is false. Consequently, Cantor's diagonalization proof demonstrates that the set of R is *uncountably infinite*. Whereas the proof illustrates this with the set of real numbers bound by the interval $(0, 1)$, this is true for the full spectrum of real numbers as well.

Besides the set of real numbers R , it can be shown that there exist other uncountably infinite sets. Using the same diagonalization method and proof by contradiction, it can be demonstrated that the assumption that $N = \delta(N)$ leads to a contradiction; more generally, no powerset of set S is the same size as set S . However, because the

Table 6 Diagonalization Proof

		Position (0, 1)					
Natural numbers		1	2	3	4	5	...
	1	.4 .5	5	7	2	8	...
	2	.2	4 5	8	0	0	...
	3	.8	0	9 0	1	7	...
	4	.9	6	0	4 2	5	...
	5	.4	7	4	5	5 6	...
	⋮	⋮	⋮	⋮	⋮	⋮	⋮

Note: Number not in correspondence with the set of natural numbers is D , where $D = .55026...$

Source: Compiled by Nathan E. Fosse.

cardinality of any powerset is 2^n , the powerset of $N = 2^{\aleph_0}$. It can also be shown that $2^{\aleph_0}$ is the same size as R , even though both sets are uncountably infinite. Because Cantor also called the continuous finite set R the *continuum*, he called the powerset of N and the size of R , both with cardinality of $2^{\aleph_0}$, the “cardinality of the continuum.”

However, other powersets are larger than the “cardinality of the continuum.” For example, the powerset of real numbers in $(0, 1)$ will not be in correspondence with $(0, 1)$. Because the number of elements in $(0, 1)$ is $2^{\aleph_0}$, the powerset of $(0, 1)$ is $2^{2^{\aleph_0}}$. This, by definition, is larger. Even higher powersets are possible, and none can be put into a one-to-one correspondence with each other.

In summary, this section has shown that Cantor’s discoveries of set theory in the late 19th century developed a new, formal method for studying the properties of infinity. Cantor proved the existence of two types of infinity: (1) countably infinite sets of cardinality $\aleph_0$, equal to the set of natural numbers N ; (2) uncountably infinite sets of cardinalities, which are equal in size to the set of real numbers R , or greater, ad infinitum. One important question is whether there exist any infinite sets of intermediate size, greater than the cardinality of N and less than the cardinality of R (the cardinality of the continuum). Cantor’s famous *continuum hypothesis* is that there is no infinite set between the cardinality $\aleph_0$ and the cardinality of the continuum (which, if true, would mean that the continuum could be expressed as $\aleph_1$). However, later mathematicians, notably Kurt Gödel and Paul

Cohen, established that the continuum hypothesis could not be proven or disproven given existing axioms; mathematically, it is *undecidable*.

Set Theory as a Foundation of Mathematics

Following Cantor, the development of set theory was largely informal, without a system of axioms to clarify assumptions are adopted when applying set theory. The major flaw with Cantor’s informal or *naïve* set theory is that it is *inconsistent*; that is, it contains paradoxes. Thus, by the early 20th century, it was clear that set theory required a system of axioms to precisely state the basic assumptions required. The most obvious reason for developing axiomatic set theory was to avoid paradoxes identified paradoxes when informal set theory is self-referential. The most famous is *Bertrand Russell’s paradox*, dealing with *universal sets* (or the “sets of all sets”). Because universal sets by definition include everything, including themselves, a universal set that excludes itself is a member of itself if and only if it is not a member of itself. In formal terms: $U = \{x \mid x \notin x\}$. The question then is: Does U contain itself?

To avoid this paradox, as well as to clarify other features of set theory, Ernst Zermelo and Abraham Fraenkel developed an axiomatic system of set theory that is widely referred to as *Zermelo–Fraenkel theory* or *ZF theory*. A more controversial variant of this axiomatic theory incorporates the Axiom of Choice, which, for subtle reasons, is included to address concerns regarding properties of infinite sets. ZFC, accordingly, refers to ZF theory with the Axiom of Choice. Despite controversy regarding ZFC, ZF theory is the most widely accepted axiomatic set theory. Based on eight axioms, ZF set theory can show that the majority of modern mathematics can be expressed as sets.

Set theory is still largely considered to be the foundation of modern mathematics. However, as Gödel proved with his famous *incompleteness theorem* in 1931, any logical/mathematical system that relies on axioms cannot be both consistent and complete. Moreover, in practice, as noted by Stephen Pollard, set theory has been only a patchwork theory, falling short of its goal as a grand, unifying theory of the mathematical universe. Despite its limited success as a metamathematical theory, why has axiomatic set theory remained as

the most widely accepted means of integrating mathematics?

There are four reasons for the widespread, continued adoption of set theory as a viable theoretical foundation. First, set theory remains useful for integrating and unifying disparate subfields of mathematics. Since the late 19th century, mathematics has specialized tremendously, from about 40 areas of specialty to at least 3,000 or more areas of specialty in the third decade of the 21st century. Consequently, set theory is useful for communicating and unifying research as the field of mathematics becomes increasingly specialized.

Second, set theory is widely entrenched in mathematics because of path dependency: Put simply, generations of professional mathematicians have been educated in set theory and socialized into adopting its theoretical paradigm. Despite Cantor's controversial findings, modern set theory is relatively uncontroversial. Third, although not complete, set theory is still relatively comprehensive and useful enough to integrate and define basic concepts. Fourth, as of the time of writing, axiomatic set theory has not resulted in any identifiable paradoxes. Finally, notwithstanding many gaps in set theory, no other competing theory has been widely accepted enough to be sufficient to address present knowledge gaps any better than set theory.

It remains to be seen whether set theory will ever be the unifying foundation for all of mathematics. There are undoubtedly supporters of axiomatic set theory. Despite the utility of axiomatic set theory, it has yet to be determined whether ZF set theory is consistent. So far, ZF set theory has not resulted in any inconsistencies despite its development in the early 20th century. On the other hand, as Gödel's incompleteness theorem states, the consistency of ZF theory will never be provable using standard mathematics. In short, there may never be any way to determine whether ZF set theory is consistent.

Conclusion

Set theory has shown great promise for integrating many aspects of mathematics and for developing a formal language that incorporates logic and mathematics. However, set theory raises important questions regarding the nature of knowledge: Are there really such entities as infinite sets despite

the fact that they do not exist in our observable reality? In short, are we discovering anything real with set theory, or are we *inventing* concepts and their relations in an elaborate formal language? Moreover, debate continues regarding the application of axiomatic set theory. Specifically, why should we accept some axioms as self-evident while rejecting others? As Noson Yanofsky points out, there are no simple answers to these questions; more work remains to be done.

Nathan E. Fosse

See also Complex Systems; Logical Theory; Statistics

Further Readings

- Dauben, J. W. (1983). Georg Cantor and the origins of transfinite set theory. *Scientific American*, 248(6), 122–131.
- Ferreirós, J. 2007. *Labyrinth of thought: A history of set theory and its role in modern mathematics* (2nd ed.). Springer.
- Gullberg, J. (1997). *Mathematics: From the birth of numbers*. W. W. Norton.
- Halmos, P. R. (1974). *Naive set theory*. Springer.
- Pollard, S. (1990). *Philosophical introduction to set theory*. University of Notre Dame Press.
- Yanofsky, N. S. (2013). *The outer limits of reason: What science, mathematics, and logic cannot tell us*. MIT Press.

SOCIAL CONSTRUCTION OF SCIENTIFIC KNOWLEDGE

There is no one single approach to the study of the social construction of scientific knowledge. The more conservative orientations analyze the institutional, professional, and normative aspects of science while still implicitly maintaining the realist assumptions of the disciplines they investigate. More radical forms view objective facts as constructed and hence contingent on a number of social factors, that if prearranged differently, may result in alternative scientific outcomes and knowledge. Despite this diversity, the organization, professionalization, and overall practice of science, like all human endeavors, are inevitably social. The work of scientists and the subsequent knowledge

their activities produce contain a combination of interpretive processes similar to the mundane social practices of everyday life. But because science is equated with reason, progress, and truth, and, relative to other institutions, enjoys unrivaled authority, many tend to overlook its social context. It was not until the mid-20th century that science became an object of sociological study.

Acknowledging that scientific knowledge is socially constructed does not necessarily imply that it is not a reflection of reality. What it does suggest is that people have to come together to interpret, negotiate, and communicate with one another to make the findings of science possible, meaningful, and legitimate. Moreover, the naïve and taken-for-granted realism and reification embedded within science motivated many sociologists, philosophers, and historians to problematize its reputation as constituting pure fact, objectivity, and universality. This mandate is not without controversy, as it has been perceived as attacking the work of science and undermining the very foundations from which its knowledge is established and made credible.

Robert Merton and Institutional Imperatives

The institutional approach to the study of scientific knowledge dates back to the 1940s with the pioneering work of Robert Merton. Rather than problematizing the objectivity and unity of science, he was more interested in understanding the social forces and institutional factors that shape an ideal form of science. Merton described the universal norms and values that allow for science to function properly. He listed four major institutional imperatives: *universalism*, *communism*, *disinterestedness*, and *organized skepticism*. Merton described universalism as the institution's capacity to evaluate scientific output independent of the individual characteristics of those who contribute findings. Gender, race, and religion do not play any role in the system's ability to reward scientific achievement. Communism represents the accumulation of scientific knowledge as a collective achievement that is to be shared freely, and not hoarded or protected as individual property. Disinterestedness reflects each scientist's concern for overall knowledge at the expense of their own self-interest and personal promotion; truth is prized over all other motivations. Finally, organized

skepticism suggests that scientists scrutinize their own and their colleagues' results until consensus in the community is achieved. When researchers deviate from these imperatives, Merton believed that the system would correct itself through sanctions. This institutional approach explores the professional and normative aspects of science while still implicitly maintaining realist assumptions. Although science as an institutional organization is investigated, its truth claims are taken for granted. If anything, the normative structure of science accounts for how and why this institution is excellent at providing truth and eliminating the potential for personal, social, and political biases.

The Ambivalence of Science

In 1974, Ian Mitroff critiqued and expanded upon Merton's idealized vision of science with his own *counter-norms* of particularism, individualism, interestedness, and organized dogmatism. Mitroff found these norms by talking with a number of moon scientists who were debating the advent of new data, which might support or hinder their own and their rivals' hypotheses. The scientists in Mitroff's sample were not always judged by their peers independently of their social attributes, since the professional reputations of scientists would go a long way in how people would value their contributions and trust their findings. He also showed that researchers are not a unified collective with their results belonging to the community as a whole, committed to a primary goal dedicated to scientific knowledge and progress. Rather, scientists can be quite protective of their own discoveries and often seek their own self-interest, motivated by personal success and recognition. Finally, Mitroff found that scientists are not solely committed to a balanced approach of critically evaluating their own work and that of their peers, but instead tend to view their own discoveries as more objective and valuable than others. Indeed, scientists had a great stake in proving their own theories as superior to others, and many passionately fought to realize their version of the truth over their rivals. In this way, the scientists' pet scientific theories would very much structure their approach to interpreting the data at hand. Mitroff's case study uncovered the ambivalence inherent in scientific research, as it would often oscillate between norms and counter-norms.

Both Merton and Mitroff studied the social organization of science and the attitudes and expectations researchers have of themselves and each other. They both provided normative explanations of how science manages to function. The difference in their work appears to be how science organizes itself. Is it a collective venture with norms providing the necessary conformity and discipline for trusted knowledge to emerge, or are their *counter-norms* acting as well, that are individually negotiated through the fact that scientists cannot help but be human, passionate about ideas, and self-interested? These contrasting motives represent a continuum of expression. Scientific social worlds, like any other institution, operate according to both collectivist and individualist logics. Although their analysis was sociological, Merton and Mitroff did not directly address the social construction of scientific knowledge. This neglect resulted in a series of critiques, with many accusing them of functional explanations, conflating traditional scientific ideology with actual research practice. Scholars would turn away from these schools of thought and find inspiration in the work of Thomas S. Kuhn, who considered how scientific knowledge actually accumulates and radically changes through history, rather than how it ought to work in ideal normative conditions.

Thought Collectives and Paradigms in the Construction of Scientific Knowledge: Fleck and Kuhn

The premise that scientific knowledge is socially constructed took shape during the mid-1930s with the Polish biologist and physician Ludwik Fleck's (1896–1961) research into the subjective nature of scientific facts. His findings led him to suggest that scientific *truth* claims are problematic because they are not passively discovered but rather generated by subtle moral and social processes. He argued that established scientific knowledge is interpreted and organized around *thought collectives* that change over time. These socialized ways of observing and interpreting data are discriminate. Facts are not solely true or false based on their more or less accurate representation of reality but are rather much more ambiguous in that their acceptance must align with the interests and goals of a community of experts. The training

of scientists tends to generate reified *thought styles* that influence and constrain the way data are interpreted and organized. Social, political, and historical contexts play a role in how specific scientific concepts win out over others. Alternative interpretations of evidence are excluded or ignored not because they are less valid but because they lack the backing of a strong *thought collective* and threaten the favored concepts these communities employ. Fleck was one of the first researchers to demonstrate that scientific knowledge is not a direct reflection of empirical evidence but that evidence is often interpreted in ways to fit prior theoretical conceptions and assumptions.

Despite Fleck's keen analysis, his insights remained ignored for decades. It was not until the 1960s and 1970s that Fleck's work began to be taken more seriously. Thomas Kuhn (1922–1996), a physicist and historian of science, credited the work of Fleck and expanded on a number of his themes in crafting his own massively influential thesis on the social nature of scientific change. In his book, *The Structure of Scientific Revolutions*, Kuhn describes the advancement of scientific knowledge not as purely cumulative and rational but as a result of sudden and disruptive changes in belief or *revolutions*. Standardized questions, theories, methods, and standards of evidence, known as *paradigms*, are modeled after the work of exemplary scientists and then entrenched in the professional education and socialization of scientists and the implementation of scientific practices to follow. These convergent ways of knowing and doing achieve prolonged states of *normal science*, whereby scientists spend most of their time in puzzle-solving. This means that rather than seeking out new radical ideas, scientists take on faith the core theories of the paradigm and merely try to figure out how to fit the empirical world into the conceptual boxes that flow logically from the theory at hand. According to Kuhn, this practice of normal science is the conventional everyday practice of science. Similar to Fleck's thought collectives, normal science demands conformity from scientists in that their interpretation of experimental results corresponds to the preestablished general consensus provided by the ruling paradigm. Scientific research that is not supported by conventional models and deviates from the dominant paradigm is rarely successful. For Kuhn, science

lacks authentic critical self-reflection and, ironically, innovates best when scientists conform rigidly to, and doggedly adhere to, previously held paradigms of thought. Rather than allow researchers the freedom to question the implicit assumptions and biases underlying normal science, the dominant paradigm forces practitioners to conform to conventional standards. Counterevidence and anomalies threaten a paradigm and are thus heavily sanctioned or ignored by the majority of supporters of normal science who have built their careers and reputations around its basic tenets.

Paradigms as Interpretive Constraints

Both Fleck and Kuhn claim that scientists' collection of data and evaluation of evidence tend to be tightly bound to their paradigm's underlying theoretical assumptions and predetermined rules of interpretation. David Rosenhan provides a salient example of a medical paradigm's power to force conformity. His experiment questioned the taken-for-granted assumption that expert systems of medical knowledge are purely objective and demonstrated the routinization, social construction, and overall constraining force of paradigmatic inertia. Specifically, Rosenhan offered a critique of the subjective nature of psychiatric diagnoses in the modern era. He asked, "If sanity and insanity exist, how shall we know them?" His answer had a profound impact on the mental health community. The Rosenhan experiment consisted of eight research associates who, on the basis of their false claims of hearing voices they defined as *hallow* and *empty*, admitted themselves to psychiatric hospitals for the sole purpose of seeing whether medical experts could identify them as sane. On this basis alone, Rosenhan's *pseudo-patients* were misdiagnosed with schizophrenia. After they were admitted, they acted normal, were honest about their life histories with hospital staff, and informed them that the voices had ceased. Despite their sane behavior while institutionalized, nurses and psychiatrists did not view pseudo-patients as *normal*. Much like the reification tendencies inherent in Kuhn's paradigms and Fleck's thought collectives, Rosenhan found that labels associated with mental illness act as conceptual frameworks in which hospital staff interpreted patient interactions. Rosenhan's conclusion is that experts interpret

and process all patient behavior solely as indicative of a presupposed underlying pathology. The irony here is that although psychiatrists could not detect the sanity of Rosenhan's research associates, labeling them as *schizophrenia in remission* upon their discharge, actual hospitalized patients did in fact see through the ruse. Rosenhan suggested that if psychiatrists could not detect the sanity of their pseudo-patients, then the true validity of the diagnoses resided not in an individual's inherent nature but in the framework of expectations provided by the environment and organization itself. That is, the suffering of mental illness is real, but the diagnostic process is paradigm dependent and socially shaped by theoretical expectations.

Gestalt Shifts and Incommensurability

In spite of paradigmatic communities' reluctance to change, revolution is unavoidable. Scientists eventually become unsatisfied with conventional treatment of anomalous data, and as this evidence grows, the validity of their theoretical models becomes impossible to sustain. This inadequacy results in the paradigm's falling into crisis, initiating a competition among new, often incompatible paradigms as to which will move in to take the place of the prior paradigm. There is an intense antagonism between the normal science of the old paradigm, which dismisses an anomaly and defends its ground, and new paradigms that embrace the anomaly as vital to their theoretical validity. Scientists choose one of these competing paradigms, with the old guard of normal science eventually being replaced by younger revolutionary converts. Kuhn argues that theory choice is not solely based on logic, since none of the new paradigms have yet had a chance to prove their worth. Instead then, persuasion, rhetoric, and social-psychological factors play a major role in influencing what emerging paradigm wins out. Once this new paradigm is established and is seen as superior to that which came before, Kuhn describes a *paradigm shift*, which represents a gestalt shift in the scientific community. A new paradigm allows scientists to view evidence differently and make sense of anomalies within a new epistemological context. This is the revolutionary nature of scientific change. These rare gestalt shifts have scientists reconfiguring the way

evidence is interpreted, which are counter to the explanations provided by previous paradigms. Kuhn believed that new paradigms are always incommensurable with those that came before. After a scientific revolution, Kuhn suggests that scientists are no longer able to explain current findings under the old paradigm. Evidence is seen in a completely new way. The nature of scientific problems and their solutions are redefined. He provides the displacement of Aristotelian science by Galileo's physics as an example of a paradigm shift and leading to incommensurable ways of seeing the world. Given the motion of a swinging stone on a string, each observer would view its motion differently. Aristotelians would view a swinging stone as a falling body whose nature and *desire* is to rest at the center of the earth, but due to the string, it can only fall with difficulty. Galileo would interpret the same swinging body as a pendulum, which almost achieves an infinite periodic cycle. This is an excellent example of the gestalt shift that accompanies paradigm change; the same physical phenomenon is seen according to an entirely different system of logic and interpretation. It was not a matter of the new paradigm being a better reflection of reality, but rather it constituted a whole new way of viewing reality that was inconceivable prior to Galileo.

Revolutionary paradigms eventually settle and become established, thus transforming into normal science. This period has its own new generation of researchers who conform to the convergent set of expectations of what makes for good science and true theories. The previous crisis and subsequent revolution are made invisible in scientific textbooks, while it is made to look as though the new paradigm were an inevitable and conservative progression from the old rather than a radical and discontinuous shift, as the new paradigm settles into business as usual. Yet Kuhn's incommensurability thesis problematized the popular view of science as a cumulative and always progressive march toward a greater stock of objective truth about the world. Accentuating the subjective, cyclical, and collective nature of scientific change, the notion of scientific revolutions, and the incommensurability between paradigms that would result became the most contested and influential aspect of Kuhn's work.

Seven years after the original publication of *Structure*, Kuhn answered his critics and clarified

his incommensurability thesis. Specifically addressing the translation problems thought to exist between paradigms, Kuhn acknowledged that scientists have the potential to see different viewpoints without experiencing the gestalt shift that he once described as tantamount to conversion. This process is possible as long as each discipline refrains from the dogmatic bias that tends to contextualize the other's foundational axioms as trivial anomalies relative to their own favored paradigm. Striving toward cooperation and not necessarily consensus, they can each come to understand contrasting theories in their own terms while being sensitive to alternative interpretations. Kuhn admits that this hermeneutic exercise is essentially how historians of science conduct their own research.

Trading Zones and Boundary Objects

Peter Galison's *trading zone* metaphor extends Kuhn's insights regarding the role knowledge translation plays in transcending paradigmatic barriers. These zones are characterized by what Galison calls interlanguages and boundary objects. The construction of a localized pidgin or creole along with the identification and maintenance of shared technical-social objects, which despite having different meanings accorded to each group, act as a symbolic bridge that links the parties together for a limited time so that ideas and resources can be traded without compromising the integrity of their prescribed paradigms. Many multidisciplinary case studies have since demonstrated how two or more different scientific subcultures with competing interests can exchange information and translate their perspectives across disciplinary boundaries without experiencing a gestalt shift or the threat of conversion. For Adele Clarke and Joan Fujimura, communities of science that may be divided by different paradigmatic logics of theory and practice can cooperate by the use of common and relevant boundary objects that help to unite otherwise divided professional and scientific social worlds. Furthermore, Susan Leigh Star and James Griesemer argue that boundary objects not only build common ground across scientific subcultures but can also facilitate knowledge translation between scientific experts and intersecting nonscientific social worlds. Benjamin Kelly importantly points out that such boundary

agreements are not always entirely fair and often become structured depending on power imbalances and inequalities between the participants.

Although Kuhn's account of science is more descriptive and relativist than Merton's prescriptive treatment, many have accused Kuhn of being too conservative in his understanding of scientific legitimacy and change. Pierre Bourdieu (1930–2002), for example, would emphasize the importance of unequal scientific capital in determining scientific change and the restructuring of academic fields. Kuhn's work nevertheless provided the foundation for what is known today as *Science and Technology Studies*. Kuhn inspired a whole generation of social constructionists who set out to problematize objective truth claims and begin to analyze the human side of how scientific knowledge is produced.

The Strong Programme and Laboratory Ethnographies

The social philosophy of science laid down by Kuhn opened the way for historians, sociologists, and anthropologists to begin to conceive of science as a rich new area to study the sociology of knowledge. Up until that time, following the work of Karl Mannheim (1893–1947), sociologists tended to treat scientific knowledge as an exception to other modes of knowledge, such as religion and politics, as a potential subject for ideological deconstruction. Mannheim followed Karl Marx in critiquing the content of knowledge in societies as ideological and distorted but always assumed that scientific knowledge was immune from these influences. This exceptional character of scientific knowledge was based on the notion that science rested on rigorous methods, such that it was beyond the level of simple claims from authority. In science, all claims had to be subjected to the careful scrutiny of systematic empirical tests. This is what Popper might call the role of *crucial experiments*, used to test whether or not given claims ought to be believed, by continued attempts at falsification through intersubjectively shared empirical experiments. But if Fleck and Kuhn were correct, that observation and experiment are themselves shaped by prior theoretical beliefs, which in turn are based on paradigms of theoretical and methodological convergence around long-held scientific cultural histories, perhaps the knowledge scientists develop in the laboratory is also open to sociological analysis. That is, like all

forms of cultural knowledge, scientific knowledge too might be socially constructed.

Edinburgh's Strong Programme

In the 1970s, several sociologists inspired by Kuhn's work launched a new campaign to try and study the socially constructed nature of scientific knowledge. David Bloor's *strong programme*, originating from the Edinburgh school in Scotland, was most famous for developing a kind of manifesto for studying the socially constructed aspects of science. The strong programme adhered to four central tenets: (1) It would be causal in that it would focus on how social conditions and processes led to the development of scientific knowledge; (2) it would be impartial with respect to truth and falsity, as both true and false beliefs equally require sociological explanation; (3) it would be symmetrical in its style of explanation. In other words, there would not be a unique method to investigate beliefs deemed to be true versus those deemed to be false; and (4) it would be reflexive, such that in principle its patterns of explanation about scientific knowledge could also be turned around and applied to social-scientific knowledge as well. This school of thought became quite controversial in that scientific knowledge was posited to be relative to social interests, structures, and processes; indeed, its final thesis of reflexivity meant that even sociological knowledge was now seen to be on shaky ground since it also had to be seen as a cultural construct over and above a direct correspondence with reality. A related school of thought to the strong programme was that of *interests theory*, put forth by Barry Barnes. This was an attempt to explain the content of scientific knowledge that was generated by certain groups based on macro-sociological variables such as social class, gender, and political affiliation. These variables were applied by Barnes to consider why, for example, certain theories of evolution, brain chemistry, or Mendelian genetics might have been shaped by the ideological interests of the scientists at work in these times.

Ethnographic Laboratory Studies and the Experimenter's Regress

As the strong programme was starting to make waves, the sociologist Harry Collins pioneered the tradition of ethnographic laboratory studies. He

began to visit scientific laboratories across England with the hope of investigating firsthand how scientific knowledge might be socially shaped. Studying the replication of lasers, Collins showed that laboratory technologies such as the lasers were not easily built. In fact, it took months and even years for some scientific groups he studied to correctly assemble and finally fire the laser. This is because constructing lasers took on a craftlike character and required a good amount of tacit knowledge over and above purely codified forms of knowledge. This finding was a direct challenge to Popper's assumption that experiments can easily be replicated in order to try and falsify tentatively held conjectures, since replication itself was not always easy or even possible. This point became even more evident as he began to study the scientific controversy surrounding the existence of gravity waves. Although everyone can agree when a laser has fired correctly, it is more controversial to know whether or not a gravity wave detector is functioning correctly. This is because the scientific community was split as to whether these theoretically exist and can in principle be detected from Earth. Does a true gravity wave detector show signs of the phenomena of interest, or should it, if correctly built, show nothing? Collins found that judgments of whether experiments were run properly depended largely on the previous beliefs scientists carried about the theoretical existence of gravity waves in the first place. This is what Collins refers to as the *experimenter's regress*, the idea that there is no neutral place for observation to settle scientific controversies, since interpretations of what should be observed and what the results mean are always linked to prior theoretical assumptions. Instead of one important *crucial experiment* to settle the controversy of gravity waves once and for all, Collins showed how social influences such as the idiosyncratic perceptions of scientists' professional credibility are more important in bringing about closure. His belief in the importance of social and micro-political processes in the shaping of scientific opinion led to his research program of *empirical relativism*. Here, he would actively seek out open scientific controversies, since what was actually true and false were still in process and up for debate between rival scientific groups. His work with Trevor Pinch to analyze scientific controversies as a way into viewing the *interpretative*

flexibility of scientific findings while observing social and political pressures impacting this became extremely influential.

Anthropological and Ethnomethodological Investigations

Bruno Latour and Steve Woolgar produced their own anthropological investigation of a science laboratory in the 1970s, taking issue with the idea that scientific facts were *discovered*, and arguing instead that they are more accurately understood as socially constructed artifacts. Analyzing the case of scientists working on the discovery of thyrotropin-releasing factor (TRF), they argued that what began as tentative claims about the potential existence of TRF became more and more *fact-like* and objective as it became applied and circulated through the spread of scientific publications. Thus, it was not further experiment that led to more certain claims about the existence of TRF but rather the social process of *black-boxing*, the once tentative conjecture into a stabilized fact. What started as a local uncertainty was transformed into a global certainty through the process of scientific communication, serving to reify the belief.

Karin Knorr-Cetina also entered into laboratory sites using an ethnographic approach in the early 1980s, arguing for the artificiality and decision-impregnated aspects of scientific experiments, and calling into question any notion of true neutral observation unshaped by previous theoretical beliefs. Theoretical beliefs could be seen as bound up in the technologies of the laboratory. Latour and Woolgar referred to these technologies as *inscription devices*, since algorithms were used to generate numeric print-outs instead of enabling direct observations of sensory phenomena. Knorr-Cetina would show the *indexicality* of laboratory processes, in that research decisions are not made entirely on rational grounds but also are subject to the availability of equipment and the methodological and practical norms at play in the local environment. Later, she would study how different laboratories and sciences operated according to different standards of practice, or what she called different *epistemic cultures*. Sharon Traweek would likewise study the cultural differences between how scientific evidence in high-energy physics was dealt with in Japan versus the United States. Michael Lynch would utilize an ethnomethodological approach to

consider how local contingencies shape scientific practices, how the verbal nuances of shop-talk reveal processes of uncertainty, and how observation and experiments are often manipulated and controlled in ways to present certain versions of reality over others. Many other studies of laboratories followed, all trying to argue for the sway of paradigmatic beliefs, theoretical motivations, social and political contingencies, and institutional cultures that led to the social construction of scientific knowledge in varying ways.

Incorporating Nonhuman Actants in the Social Construction of Scientific Knowledge

Latour eventually came to a different conception of the relevance of laboratory studies by the early 1990s. He argued that while the goals of these studies were to show how scientific knowledge is socially constructed, they actually showed there was much more happening “on the bench” than the mere realization of previously held social and theoretical interests. Latour argued that not only do scientists have agency in shaping the construction of scientific knowledge but so also do a range of nonhuman actors in the laboratory, which often act against the interests of the scientists involved. Thus, in Steven Shapin and Simon Shaffer’s study of the debate between Thomas Hobbes and Robert Boyle in the 1660s about the existence of a vacuum, public experiments would be run with an air pump, and the settlement of these scientific controversies would often rest in the hands of what Latour called *nonhuman actants* (in this case mice suffocating, or candles going out), that were enrolled in the experimental setting to convince the public of one side of the debate. Andrew Pickering would similarly argue that a close ethnographic examination of experiments reveals what he calls the “dance of agency” between human and nonhuman actors as they exchange active and passive roles through the process of observational work through time. Latour would take this as a launching pad to critique social constructivists, whom he accused of the modernist error of trying to purify the world of human ideas from the world of nature, and privileging the former in explanation of the latter. If we could only retie the production of knowledge to what he refers to as the *middle-kingdom*, the place where the complex processes of mediation between

human and nonhuman actors happen in a common world, we would have a much better understanding of how knowledge gets made in the messy, practical contexts of “science in action.”

Criticisms of Social Constructionist Perspectives

The constructivist sociology of science has been criticized for being gender blind. Sandra Harding critiques mainstream science (and the sociology of science) as masculinist and Eurocentric and proposes instead a new successor science based on her belief in standpoint theory. She argues that groups who are marginalized in society have a better grounding for, and interest in, generating truth claims that are undistorted by the biases of the dominant and often patriarchal status quo. Her work has inspired many, such as Emily Martin’s study of the problematic masculine assumptions that lie behind the science of human reproduction, and Donna Haraway’s analysis of how constructions of primate society reflect certain patriarchal definitions of the family. Donna Haraway would depart from Harding’s strong version of standpoint theory, however, and argue that all positions, whether marginal or not, are limited by their location that permits only a partial perspective on reality. Pierre Bourdieu was also critical of traditional constructionist accounts in that they would miss the fact that the laboratory sites studied were embedded in wider institutional power struggles between competing actors in the structured hierarchy of science. Stephan Fuchs provides a similar argument that scientists’ location on the institutional fringe, frontier, or center of scientific networks will largely determine the type of research questions (e.g., in Kuhn’s language, either normal or revolutionary) are pursued. Thus, a better institutional, contextual, and macro-sociological approach to the sociology of scientific knowledge would help to rebalance the field by providing a better understanding of the myriad forces that affect scientific knowledge production across the range of settings in which it is realized.

Controversy and Future Directions

In 1996, Alan Sokal, a professor of physics at New York University, published an article entitled “Transgressing the Boundaries: Toward a

Transgressive Hermeneutics of Quantum Gravity” in the postmodern journal *Social Text*. Sokal combined a number of ideas from quantum physics with a variety of postmodern concepts to demonstrate the relativist and subjective nature of reality. After its publication, Sokal confessed that his writing was not an authentic contribution to a postmodern view of reality but a nonsensical mishmash of sophisticated terms woven together illogically. In other words, it was announced as a hoax. His practical joke, which became known as the *Sokal Hoax*, was designed to demonstrate the absurdity of postmodern theory, and further, of the social constructionist perspective on science. His goal was not only to reassert the validity of realism and the scientific method in the hopes of revealing the anti-scientific agenda of social constructionists but also to awaken the Left from its nonsensical postmodern vision of progress. This was a direction, Sokal believed, that was undermining democracy and positive political change. Many STS scholars took offense and attempted to correct what they believed was Sokal’s misguided assessment of their work. The heated debates that followed were known as the climax of the science wars, which had been brewing for years between traditional realist philosophers, natural scientists, and this new emerging relativist school of social constructionism from the 1970s. With no clear winners, one unintended consequence of the science wars was the political Right’s ability to promote anti-intellectualism. Philosopher James Robert Brown argues that the Sokal Hoax uncovered something more important than the legitimacy of the social construction of science. He suggests that it exposed the political undercurrents behind attitudes toward science and argues that the science wars and Sokal’s hoax informed the public that the antagonistic or favorable views toward orthodox science are much more complicated than the common but false “anti-science Left versus pro-science Right” dichotomy.

Since these debates, the sociology of scientific knowledge has begun to expand its range and scope. Following Latour’s earlier advice to “follow scientists around society,” recent scholars have examined how science and experimental work play out and have relevance for an array of topics as wide as the social and cognitive sciences, education, medicine, genetics, politics, forensics and crime, scientific governance, the environment, economics,

law, the public participation of science, as well as multicultural and global issues of knowledge production. What has remained is a dedication to detailed historical and ethnographic case studies to uncover the micro-sociological practices that show how local contingencies, politics, and negotiations are crucial to understand how knowledge is sedimented, universalized, and spread. This approach of local, fine-grained ethnographic analysis has also extended outward into questions concerning technology, its relation to science, and its incorporation and reshaping in relation to everyday users across a range of social contexts. What began as a trail-blazing exercise in trying to push for radical and relativistic epistemologies of science from the 1970s has morphed into a well-institutionalized and respected approach to the study of knowledge production, with a range of interdisciplinary applications. The sociology of science remains a field ripe for investigating the structuring role of theory in knowledge production and how such theories are always richly embedded in culture.

Benjamin W. Kelly and Antony J. Puddephatt

See also Paradigm; Phenomenalism; Philosophy of Science; Scientific Revolutions

Further Readings

- Barnes, B. (1977). *Interests and the growth of knowledge*. Routledge.
- Bloor, D. (1976). *Knowledge and social imagery*. University of Chicago Press.
- Bourdieu, P. (2004). *Science of science and reflexivity*. University of Chicago Press.
- Brown, J. R. (2001). *Who rules in science: An opinionated guide to the wars*. Harvard University Press.
- Clarke, A., & Fujimura, J. (1992). *The right tools for the job: At work in the 20th century life sciences*. Princeton University Press.
- Collins, H. (1985). *Changing order: Replication and induction in scientific practice*. University of Chicago Press.
- Collins, H. (2010). *Gravity’s shadow: The search for gravitational waves*. University of Chicago Press.
- Collins, H., & Pinch, T. (1993). *The Golem: What you should know about science*. Cambridge University Press.
- Fleck, L. (1979). *Genesis and development of scientific fact*. University of Chicago Press. (Original work published 1935)
- Fuchs, S. (1992). *The professional quest for truth: A social theory of science and knowledge*. SUNY Press.

- Galison, P. (1997). *Image & logic: A material culture of microphysics*. University of Chicago Press.
- Hackett, E., Amsterdamska, O., Lynch, M., & Wacqman, J. (2008). *The handbook of science and technology studies* (3rd ed.). MIT Press.
- Haraway, D. (1989). *Primate visions: Gender, race and nature in the world of modern science*. Routledge.
- Haraway, D. (2001). Situated knowledges: The science question in feminism and the privilege of partial perspective. In M. Lederman & I. Bartsch (Eds.), *The gender and science reader* (pp. 169–188). Routledge.
- Harding, S. (1986). *The science question in feminism*. Cornell University Press.
- Kelly, B. (2017). The social psychology of compromised negotiations: The emergence of asymmetrical boundary objects between science and industry. In B. Brewster & A. Puddephatt (Eds.), *Microsociological perspectives in environmental sociology* (pp. 190–203). Routledge.
- Knorr-Cetina, K. (1981). *The manufacture of knowledge: An essay on the constructivist and contextual nature of science*. Pergamon Press.
- Knorr-Cetina, K. (1999). *Epistemic cultures: How the sciences make knowledge*. Harvard University Press.
- Kuhn, T. S. (1962). *The structure of scientific revolutions*. University of Chicago Press.
- Latour, B. (1993). *We have never been modern*. Harvard University Press.
- Latour, B. (2007). *Reassembling the social: An introduction to actor-network theory*. Oxford University Press.
- Lynch, M. (1993). *Scientific practice and ordinary action: Ethnomethodology and social studies of science*. Cambridge University Press.
- Mannheim, K. (1936). *Ideology and utopia*. Routledge.
- Martin, E. (1987). *The woman in the body: A cultural analysis of reproduction*. Beacon Press.
- Merton, R. (1973). *The sociology of science: Theoretical and empirical investigations*. University of Chicago Press.
- Mitroff, I. (1974). Norms and Counter-norms in a select group of the Apollo moon scientists: A case study of the ambivalence of scientists. *American Sociological Review*, 39, 579–595.
- Pickering, A. (1995). *The mangle of practice: Time, agency, and science*. University of Chicago Press.
- Popper, K. (1959). *The logic of scientific discovery*. Hutchinson and Company.
- Rosenhan, D. (1973). On being sane in insane places. *Science*, 179(4070), 250–258.
- Shapin, S., & Shaffer, S. (1985). *Leviathan and the air pump: Hobbes, Boyle, and the experimental life*. Princeton University Press.

- Sokal, A., & Bricmont, J. (1998). *Fashionable nonsense: Postmodern intellectuals' abuse of science*. Picador.
- Star, S. L., & Griesemer, J. (1989). Institutional ecology, 'translations' and boundary objects: Amateurs and professionals in Berkeley's museum of vertebrate zoology, 1907–1939. *Social Studies of Science*, 19, 3. <https://doi.org/10.1177%2F030631289019003001>
- Traweek, S. (1988). *Beamtimes and lifetimes: The world of high energy physicists*. Harvard University Press.

SOFTWARE ENGINEERING

Since the terms *software* and *software engineering* (SE) began to be used, they have undergone significant changes. While the meaning of the former has evolved from a specific program to a complex product, the latter has become a professional discipline affecting almost every industry. Spectacular improvements in hardware, memory, and storage capacity, along with the new computing architectures, have paved the way for sophisticated computing systems over the past five decades. Today's world cannot be run without software, and our global society could not function without professional software systems. Software is embedded inside everything from simple consumer electronics to medical systems to command and control systems. Therefore, it provides a gateway to worldwide information networks; infrastructures, transportation, industry, and financial systems are all controlled by computer-based systems. SE has become critically important, and therefore, software engineers have responsibilities both to the engineering profession and society.

SE is an applied discipline encompassing processes, methods, tools, and principles in order to build reliable, maintainable, and large-scale software systems with high productivity and quality. The inherent constraints in SE stem from its intangibility, complexity, and human dependency, and thus, the main problem in SE is generally viewed as that of making the connection between the abstract world and the physical world. To that aim, software engineers are concerned not only with technical issues but also the conceptual and theoretical background of SE as a discipline. Theory provides a conceptual framework for organizing facts and knowledge, it helps understanding the underlying mechanisms of software systems, and finally, it facilitates communication of ideas in

SE community as well as in SE research area. From an evolutionary point of view, SE can be regarded as a young discipline; however, it has transdisciplinary and theoretical foundations that incorporate multifaceted principles in other disciplines while allowing the development and use of SE-specific theories. After defining SE and noting its key concepts of SE, the entry goes on to describe the nature of software, SE as a discipline, its theoretical and transdisciplinary foundations, and the development and use of theory in the research and the software industry.

Engineering as Concept and Profession

Engineering is an applied discipline as well as a technological and problem-solving approach in which complicated artifacts are systematically planned, designed, developed, and maintained on the basis of scientific theories and principles. As a profession, it is defined by the Accreditation Board for Engineering and Technology (ABET) as “the profession in which knowledge of the mathematical and natural sciences gained by study, experience, and practice is applied with judgment to develop ways to economically utilize the materials and forces of nature for the benefit of mankind” (ABET, 1986, n.p.). Therefore, the use and choice of appropriate tools and methods is the key to engineering when playing the roles such as research, design, development, production, testing, and maintenance during engineering processes.

The concept of engineering, which emerged from the times of manufacture-based and mass-production-oriented industrial revolution, is a methodology for achieving a common goal with a group of people working together, which could not be done by individuals technically, physically, or economically. Science accumulates and organizes the theoretical and empirical knowledge for solving technological and engineering problems. While scientists perform research to create new theoretical knowledge, engineers convert this knowledge into practices, applications, devices, and processes for the design and development of products. In other words, scientists seek to know and explore what is, and engineers find out how to do. As a result, engineers perform, calibrate, and validate measurements; make approximations based on experience and empirical data; and work quantitatively whenever possible.

Software: Definition and Function

The term *software* is credited to John W. Tukey, a statistician and chemist, who first used it in print in 1958. SE, as a concept, was initially used in a North Atlantic Treaty Organization conference in 1968. Software can be defined as the computer programs that are made up of instructions running on different types of computer systems and devices, providing users with desired functions and performances and enabling users to process and manipulate data. Software is logical rather than physical, and therefore, it has substantially different characteristics than those of hardware. It is not only another word for computer programs, it also includes associated documentation and configurations required to make a program work properly. A *computer program* is the code created by a programming language or the algorithm plus data; *software*, however, is an integration of the program code with design and supporting documentations, intermediate work products including system specification and design, decisions, code configuration, test cases, maintenance, and user manuals.

By playing a dual role, software can be a product or it can be a means to deliver a product acting as the basis for the control and creation of tools and environments. Software can be generic or customized. The former is the stand-alone system developed by an organization and sold on the open market, while the latter is delegated by a particular customer to a software contractor and then developed specifically for that customer. It is also possible to categorize software into the groups such as system software, application software, product line software, scientific software, embedded software, web application, and so on, each in turn having a unique context. From the viewpoint of a software engineer, a software product is the combination of programs, data, and other computer software. In addition, this product can be viewed by a software user as a means to create information somehow making their life better.

Software Engineering as a Discipline

Engineering approaches and common engineering principles form an integral part of the theoretical and practical foundations of SE. In conventional engineering, the common approach is to move from abstract to concrete, and the final product is usually the realization of an abstract design

physically. In SE, however, as explained by Yingxu Wang, this approach is reversed. Moving from the physical world to the abstract, the final product is the virtualization and coding of a software design that expresses and represents a real-world problem. As an engineering discipline applying theories, processes, methods, and tools to build high-quality software, SE is concerned with all aspects of software production, from the phases of determining requirements and meeting system specification to maintenance of the system after delivery. SE is the use and establishment of SE principles in order to obtain reliable, measurable, and efficiently working large-scale software with high productivity. Software engineers apply these principles and look for solutions to problems within limited resources and constraints, such as time, scope, and budget.

As defined by the Institute of Electrical and Electronics Engineers, SE is “the application of a systematic, disciplined, quantifiable approach to the development, operation, and maintenance of software” (CAST, n.d.). Therefore, engineering, computer engineering, computer science, mathematics, systems engineering, management, economics, cognitive science, information science, project management, and quality management are among the major related disciplines. Accordingly, software engineers should have the knowledge of material from these disciplines to some extent. This domain knowledge can also be subdivided into knowledge areas such as SE process models and methods, software quality, requirements, design, construction, testing, maintenance, and configuration management processes. An experienced software engineer should have the fundamental knowledge of SE principles, acquire problem-solving skills, know how to govern software projects, implement software processes, choose software development tools and environments, and master computer programming as well as required programming languages.

Software Engineering Process

An SE process, which is the combination of a set of interrelated activities that are functionally coherent and reusable for SE, transforms inputs into outputs using SE tools and techniques. SE processes and activities (planning, design, construction, testing, configuration management, and

maintenance) occur both at the organizational and project levels to ensure that the software product is delivered effectively and efficiently for the benefit of all stakeholders. Therefore, SE methods and the use of structured modeling techniques provide a systematic approach to both problem-solving and software development. Software processes that are used to specify and transform requirements into a deliverable software product are included in the *software development life cycle* (SDLC) and can be linear and iterative, whereas a *software product life cycle* includes an SDLC plus the additional processes of maintenance, support, and evolution.

It is possible to classify software process models as *prescriptive* or *agile* process models. The focus of prescriptive models is the detailed identification, definition, and application of process activities and tasks, while agile models emphasize a more informal and flexible approach to software processes. It is important to note that there is no ideal or single best software process or a set of software processes. Therefore, a process adopted for a specific software project might be considerably different from a process adopted for another software project. The best of SE processes and models have to be selected, adapted, combined, and applied for each project in an organizational context.

SDLC begins with the *software requirement process*, which includes elicitation, analysis, specification, validation, and management of the needs and constraints placed on a software product and its development. The *software design process* consists of architectural and detailed design, which in turn describes how software is organized into components and desired behaviors in sufficient detail. As one of the life cycle stages, the design process produces a description of the software's structure with a set of models and artifacts serving as the basis for its construction. Software design and architecture methods, such as *object-oriented design* and *structured design*, provide a common framework for software engineers. *Software construction process* refers to the creation of working software through a combination of coding, testing, debugging, verification, and integration activities. This stage produces source files, test cases, documentation, and configuration items that have to be managed in a software project. *Software testing* is the

verification process to determine whether a program or module meets different types of expectations and provides the behaviors on a set of test cases. *Configuration management* includes systematical control of the changes to software as set forth in technical documentation, which also aims to ensure the maintenance of integrity and traceability of the software. Finally, *software maintenance* is the modification of existing software while preserving its integrity. It is performed in order to correct faults, implement enhancements, and adapt software to different hardware and software systems or environments.

Software engineers also must understand quality-related concepts, characteristics, and values, and their application to SE processes. Because they are closely interrelated, the activities pertaining to SE process quality, product quality, and software quality management have a direct impact on the quality of the software process and final product. Therefore, one of the primary responsibilities of software engineers and project managers is to balance the sometimes conflicting demands of project scope, time, cost, risk, software process, and product quality.

Theoretical and Transdisciplinary Foundations of Software Engineering

SE has usually been perceived as a part of computer science, providing basic computing theories and programming methodologies. However, the transdisciplinary foundations of SE indicate that it requires not only domain knowledge but also understanding the theoretical essences that have close relationships with other disciplines. Therefore, theoretical and empirical SE explores models, architectures, and methodologies for large-scale software development as well as the nature and mechanisms of software behaviors and the laws behind them. The main problem in SE can be seen in making connections between the abstract world and the physical world. The basic constraints of SE mostly stem from its intangibility, variability, complexity, and human dependency. Therefore, SE theories and methodologies are developed and adopted primarily for dealing with these challenges.

From an evolutionary point of view, SE can be regarded as a young discipline; however, its theoretical foundations incorporate multifaceted

principles, transdisciplinary theories, and the knowledge acquired from both theoretical and empirical studies, as well as industrial practices. For example, SE emerged from computing theory, which is thus one of its most important foundations. The fundamental models in computing can be classified into *program modeling*, *data and object modeling*, *operational modeling*, and *process and resource modeling*. Computer science provides the computational methods, computer architectures and implementations, computing objects and their abstract representations, and programming methodologies.

Mathematics is of course a key means of abstraction and thus constitutes the most general form of human knowledge in science and engineering. For this reason, SE uses it to express basic notions in forming conceptual models; in the design and treatment of architecture; and to define abstract objects, relations, and software behaviors at the highest level of abstraction. Topics such as mathematical logic, set theory, and functions and relations are essential means for modeling software architectures and behaviors. Regarding the foundations of systems science, it is also seen that the concept of systems is applied in many disciplines of science and society, and that a system may be classified into any of a variety of categories depending on its key characteristics and components. Given their highly complex and abstract nature, software systems commonly adopt system theory in their life cycle. Thus, ISO/IEC 15288 provides a process model that accepts large-scale software development as part of software systems to help in dealing with assorted problems and complexities in SE. Because system engineering is concerned with all aspects of systems development that may include process, hardware, and software, SE may be viewed as part of this more general and abstract process.

Information is the product of natural or machine intelligence, and *information science* studies the nature of information and the ways of its processing and transformation. Accordingly, the main function of languages is widely viewed as that of communicating information and expressing abstract human thoughts and behaviors. In the same context, the discipline of linguistics studies natural languages, which are oral, written, or symbolic systems for communication, thought, and self-expression. Therefore, software can be seen as

a type of behavioral and instructive information about architectural and behavioral aspects of a software system. Programming languages describe and specify this computing and instructive information about the architectural and behavioral aspects of the software system.

One discipline that has strong ties with SE is *cognitive science*, which consists of multiple research disciplines. Cognitive science studies mental processes; it explores how information is represented, processed, and transformed within the brain and in computing systems while addressing the constraints on cognitive processing of information. Software is initially generated, represented, and cognitively processed in the long-term and working memory of the brain before it can be transferred to a computer. Indeed, both problem-solving and SE have much in common in that both require high-order cognitive skills and engage a variety of cognitive components. Therefore, acquiring computer programming skills and development of large-scale software systems are based on fundamental cognitive processes. *Information processing theory* and *cognitive load theory* are examples of theories in this knowledge domain and are also extensively used in the research studies of various disciplines.

Another example for the foundational disciplines is that of *management science*, which primarily studies how an organization may be operated effectively and efficiently within given internal and external constraints and in different environments. It involves organizational theories, decision theories, operational theories, quality theories, and strategic planning. Management is a process whose main functions—planning, organizing, and controlling—are instrumental to achieve goals that may not be possible for individuals acting alone. When tracing the history of SE, it can be seen that many of its important concepts, such as requirement analysis, design testing, and quality control, were borrowed or adopted from the methods and practices developed in management science and in other engineering disciplines. Therefore, in addition to computer programming and technical aspects of software development, SE also deals with the issues of organization and management infrastructures.

Theory Use and Development in Software Engineering

Theories are commonly viewed as a coherent set of tested propositions, which are generally regarded as correct, and able to predict or explain facts or phenomena in SE. A theory should have empirical foundations and be grounded in the systematic collection of empirical evidence, thereby maximizing its potential use by practitioners and researchers. Hence, an SE theory provides a conceptual framework for explaining observed phenomena and aids in understanding the basic concepts and underlying mechanisms of software systems and their behaviors. Because SE is an applied discipline and each SE case is unique, a theory may need local adaptations, and it is expected to be relevant, formalized, and ultimately useful for the software industry.

When supporting research studies, an SE theory helps engineers to develop and combine research programs and facilitates communication of knowledge and ideas. Within the industry, it may provide software decision makers with required input regarding the selection of a method, tool, or technology for a software project. There may be three modes of theory used in SE: using theories from other disciplines as they are, adapting theories generated in other disciplines to SE, and generating theories from scratch in SE discipline.

According to Shirley Gregor and to Jo Erskine Hannay and colleagues, there are also five types of theories that may be adapted to an SE context: (1) “analysis” theories, which include classifications, taxonomies, ontologies, and describes object of a study and “what is”; (2) “explanation” theories, which explain why something happens or why a phenomenon occurs; (3) “prediction” theories, used to predict what would happen using probabilistic and mathematical models; (4) “explanation and prediction” theories, which are empirically based; and (5) “design and action” theories, which are usually prescriptive and which describe how to do things in an SE process.

Dag I. K. Sjøberg and colleagues suggest that four main parts, such as (a) constructs, (b) propositions, (c) explanations, and (d) scope, may comprise the structure of an SE theory. The constructs are the basic components in which an SE theory is expressed and to which this theory provides a

prediction or description. Propositions are formed by the relationships indicating how the constructs interact. The explanations, which are also experimental observations of propositions, are logical reasons showing why the propositions are as specified. The circumstances under which the theory is assumed to be applicable define the scope of an SE theory. Thus, constructs and the relationships between constructs form the building blocks of SE theories, and they should be derived from or associated with four archetype classes: (1) *actor*, (2) *technology*, (3) *activity*, and (4) *software system*. The subclasses *industry*, *organization*, *team*, *individual*, and *project* may extend from the parent class *actor*. A programming language, tool, technique, method, model, and process can be the subclasses of class *technology*. Planning, analyzing, designing, developing, and modifying a software system are the subclasses of class *activity*. Finally, application domain and type of software and project subclasses may belong to the parent class *software system*. These classes and subclasses, together with condition statements, define the scope of an SE theory and for whom, where, and when the theory applies.

SE requires both empirical and theoretical research. Empirical SE studies explore, predict, and try to explain the investigated cause–effect relationships between constructs of a theory and find out what types of SE constructs, which are also derived from the archetype classes, should be used in what situations and circumstances. An SE experiment is, therefore, the primary research method directly to make comparisons and to observe the effects of measures taken to improve an SE process. The research approaches for empirical studies should be based on quantitative, qualitative, or mixed research paradigms, which may also use experiment, case study, benchmarking, and standardization methodologies with statistical analysis techniques. According to Hannay and colleagues, the roles that a theory would play in SE experiments include the following: (1) The design of an experiment for the hypotheses may be justified by a theory; (2) a theory may be used to explain observations on the cause–effect relationships after an experiment; (3) a theory may be tested directly with an experiment when additional justification is necessary; (4) an existing theory may be refined and enhanced and then it needs to be tested; (5) a theory may involve and

provide structural elements as a basis for the use of another theory.

Conclusion

Historically, language-centered computer programming has been dominant in SE. Technical innovations in SE have served from time to time as a major driving force for SE trends and practices. However, certain fundamental problems remain, and some authors have claimed that the issues may be regarded as theoretical and not only practical ones. This may be supported by studies reviewing theory use in empirical SE and reporting the lack of explicit and/or relevant theory use in SE. Indeed, there is evidence that an awareness of the theoretical problems and the importance of theory use in SE does exist. However, as noted by Hannay and associates, the majority of studies appear to provide post hoc explanations of the results, or they justify research questions rather than exemplify a theory-driven research study in SE. Therefore, there have also been attempts to indicate that SE may need more general, discipline-specific, or unified SE theories as in some other disciplines.

Murat Pasa Uysal

See also Cognitive Science; Cybernetics, 20th Century; Information Theory; Systems Science

Further Readings

- Accreditation Board for Engineering and Technology. (1986). *1985 Annual Report*.
- CAST. (n.d.). What is software engineering? Retrieved January 25, 2021, from <https://www.castsoftware.com/glossary/what-is-software-engineering-definition-types-of-basics-introduction>
- Gregor, S. (2006). The nature of theory in information systems. *MIS Quarterly*, 30(3), 491–506.
- Hannay, J. E., Sjøberg, D. I. K., & Dybå, T. (2007). A systematic review of theory use in software engineering experiments. *IEEE Transactions on Software Engineering*, 33(2), 87–107.
- International Organization for Standardization/International Electrotechnical Commission (ISO/IEC) 15288. (2008). The standard for systems and software engineering-system life cycle processes. <https://www.iso.org/obp/ui/#iso:std:iso-iec:15288:ed-2:v1:en>

- Pressman, R. S. (2010). *Software engineering: A practitioner's approach* (8th ed.). McGraw-Hill.
- Sjøberg, D. I. K., Dybå, T., Anda, B. C. D., & Hannay, J. E. (2008). Building theories in software engineering. In *Guide to advanced empirical software engineering* (pp. 312–336). Springer.
- Software Engineering Method and Theory. (2014). *Software engineering method and theory*. Retrieved December 2021 from <https://www.infoq.com/news/2010/04/semat/>
- Sommerville, I. (2011). *Software engineering* (9th ed.). Pearson Education.
- SWEBOK (Software Engineering Body of Knowledge). (2014). *Guide to the software engineering body of knowledge Version 3* (Bourque, P., Fairley, R. E., Eds.). IEEE Computer Society.
- Wang, Y. (2008). *Software engineering foundations: A software science perspective*. Auerbach Publications.

SOUNDNESS

A deductive argument is valid if, and only if, it is of a form that makes it impossible for the premises to be true and the conclusion to be false. If this is not the case, then the argument is invalid. A deductive argument is sound if, and only if, it is valid and all the premises are in fact true. If not, the deductive argument is considered unsound. This entry discusses the concept of soundness, which is perhaps the single most important concept in logic.

A sound argument is guaranteed to preserve truth. Other fields can decide which premises are true, then logic can guarantee that when true premises are used in a sound argument the conclusion must also be true. Looked at this way, only two things are worth discussing: logic, which tells us which arguments preserve truth, and everything else, which tells us which premises are worth arguing from.

Soundness in Deductive Logic

A deductive argument is an argument that is intended to be deductively valid, that is, one that provides a guarantee that the conclusion is true on the basis that the argument's premises are true. In a deductive argument, the premises must provide strong support for the conclusion, such strong

support in fact that if they are true, it will be impossible for the conclusion to be false.

This is in comparison to an *inductive* argument, which is an argument that seeks to show that its conclusion is probably, albeit not necessarily, true. An inductive argument claims only that if the premises are true, then the conclusion is likely to be true, but not necessarily so.

Deductive arguments are generally expressed using terms such as *therefore*, or *it follows that*, or *it must follow that*, and so on. Inductive arguments are introduced using terms such as *it is probable that*, or *it is likely that*, and so on.

The following are examples of deductive arguments where it is apparent that the argument aims to show that the conclusion is true if the premises are true. The first argument is valid, while the second is invalid:

1. All cats are mammals.
2. All mammals have warm blood.
3. It must therefore be the case that all cats have warm blood.

1. All cats are mammals.
2. All human beings are mammals.
3. It must therefore be the case that all cats are human beings.

The difference between these two arguments is that the first argument involves good reasoning and the second, poor reasoning. The first argument is a valid argument, and the second is invalid. What is sometimes confusing is that arguments can be valid even when they have false premises and false conclusions. For example:

1. All pilots are Buddhists.
2. All Buddhists are millionaires.
3. It must be the case that all pilots are millionaires.

The premises in the above argument are false, and even though the conclusion is false too, the argument is a valid one. That is because if the premises had been true, the conclusion would have been true as well. So, the argument is in line with the definition of a valid argument. Thus, true premises are not required to achieve validity. Logically,

valid does not entail *true*. An argument can be valid so long as if the premises are true then the conclusion must be true. An argument is considered valid if the premises and the conclusion are related in the right way, and we can recognize that even if one of the premises is in fact false that if were true the conclusion would have been true, too.

Despite an argument being valid, however, we would also like for it to be true. If an argument is valid, and additionally it is the case that the premises are true, then we have what is known as a valid argument which is sound. Validity is concerned with the structure of an argument, while soundness is the substance of the argument.

A sound argument has the following properties:

1. That all the premises are true.
2. That it is valid.

Truth, which means a correspondence with reality, is our ultimate goal when we are reasoning. Soundness is the overriding aim of deductive reasoning, not just validity. If our deductive argument is both sound and valid, we have a best-case scenario. This argument is both valid and sound:

1. All cats are mammals.
2. All mammals are animals.
3. Therefore, it must be the case that all cats are animals.

In the previous argument, where we discussed pilots and Buddhists, the argument was valid but unsound. It's worth pointing out that both invalid and valid but unsound arguments can entail true conclusions. We cannot reject an argument's conclusion by simply discovering that the given argument for the conclusion is flawed.

Logical Form

Whether the premises for an argument are true or not depends on the specific content. Most logicians, however, understand the validity or invalidity of an argument as being determined solely by its logical form. The logical form of an argument is what is left behind when we abstract

away from the specific content in the premises and the conclusion (i.e., words that name things, properties of things and their relations) to leave just those elements common to discourse and reasoning. This discourse and reasoning can be about any subject matter, for example, words such as *and*, *all*, *some*, *not*, and so on. The logical form of an argument is represented by replacing specific word content with letters used as *placeholders*, or variables.

Consider the following two examples:

1. All cats are mammals.
 2. No mammals are creatures with scales.
 3. It must be the case that no cats are creatures with scales.
-
1. All orangutans are zebras.
 2. No zebras are animals.
 3. It must be the case that no orangutans are animals.

These arguments have the same logical form:

1. All A are B.
2. No B are C.
3. It must be the case that No A are C.

Arguments with this form are valid, and because they have this form, the examples just shown are valid. However, the first argument is sound, and the second argument is unsound since its premises are false. In the following:

1. All footballs are round.
 2. The Earth is round.
 3. It must be the case that the Earth is a football.
-
1. All popes live in the Vatican.
 2. John Paul II lives in the Vatican.
 3. It must be the case that John Paul II is a pope.

These arguments also have the same form as:

1. All A's are F.
2. X is F.
3. It must be the case that X is an A.

These types of arguments are invalid, and we can easily see this in the first example. The second

example might appear like a good argument since the premises and the conclusion are true, but in fact the conclusion's truth is not guaranteed by the truth of the premise. It might have been the case that the premises were true and the conclusion false. That makes this argument invalid and entails that the argument is also unsound. Although most contemporary logicians accept that logical validity is determined entirely by form, there is some disagreement. Consider the following:

1. My table is round. Therefore, it is not square-shaped.
2. Jules is a bachelor. Therefore, he is single.

These arguments appear to have the form
X is F, therefore X is not G.

Arguments of this form are not as a rule valid. However, it seems intuitively the case that it must be impossible for the premises to be true without the conclusion being true too.

Logicians would respond to this puzzle in any of a number of ways. Some could insist, controversially, that these arguments actually contain premises that are implicit, for example, "Nothing can be both round and square shaped." Or "All bachelors are single," which, while they are in themselves necessary truths, also play a role in the form of the arguments. It could also be proposed that although (notwithstanding the addition of the additional premise) there is a necessary connection between premise and conclusion, the type of necessity involved is not a *logical* necessity, which means that the argument in the simple form cannot be considered logically valid. Finally, and especially when we look at the second example, a logician might suggest that because a *bachelor* is defined as an unmarried male, the true logical form of the argument is the following universally valid form:

X is F and not G and H.

Therefore, X is not G.

It is not as easy as first thought to determine the logical form of a statement. Statements with apparently identical grammatical form can have different logical form. For example:

Nigel is a ferocious lion.

David is a lame duck.

The two statements might look the same, but only one (1) has the form X is an A that is F. From this, we can infer (validly) that Nigel is a lion. We cannot deduce that David is a lame duck in (2). This same sentence can be used in other ways and in other contexts, for example:

The prince and princess are visiting dignitaries.

At first sight, it isn't clear what the logical form of this sentence is, whether there are dignitaries that the prince and princess are visiting in which case the sentence takes the same logical form as "The prince and princess are playing guitars," or whether the prince and princess are themselves dignitaries. In that case, the sentence would have the same logical form as "The prince and princess are quivering cowards."

Depending on the form the sentence is thought to take, the inferences are either valid or invalid.

Consider the following:

The prince and princess are visiting dignitaries.

Visiting dignitaries is tiresome.

Therefore, the prince and princess are doing something tiresome.

This statement can only be considered valid if the statement is given in the first reading.

Soundness in Formal or Symbolic Logic

Because it's so difficult to identify the logical form of an argument, along with the potential to deviate from logical form to grammatical form in everyday language, logicians tend to make use of artificial languages in which logical and grammatical form coincide. These artificial languages contain symbols similar to the ones used in mathematics to represent ordinary words such as *all*, *or*, *and*, and so on. This makes it easier to set rules that can determine whether the argument is valid or not. The study of deductive argument forms to ascertain whether they are valid or invalid is termed *formal* or *symbolic logic*.

In order to evaluate a deductive argument, we need first to ask whether the premises support the conclusion by considering the form the argument takes. If they do, the argument is valid. Next, we need to ask whether the premises are

actually true or false. If and only if the argument passes both of these tests can we say with confidence that the argument is sound. If the argument doesn't pass the tests, then it is possible that the conclusion may still be true, notwithstanding the fact that no support for its truth has been provided in the premises.

Soundness and Completeness in Mathematical Logic

The idea that mathematics is logic goes back to the philosopher and mathematician Gottfried Wilhelm Leibniz (1646–1716). But during the 19th century, the basic principles of fundamental mathematical theories were established by Richard Dedekind (1831–1916) and Giuseppe Peano (1858–1932) and the principles of logic by Gottlob Frege (1848–1925).

Soundness is one of the most important properties of mathematical logic, providing the reason for seeing a logical system as desirable. Logical systems are sound if, and only if, every formula to be proven within the system is logically valid with regard to the semantics of the system. In most cases, this necessitates the rules having the property of preserving truth.

Many proofs of soundness are trivial; for example, when we are discussing an axiomatic system, the proof of soundness is simply the validation of the axioms, and that the rules of inference preserve validity, or weaker property of truth. Most axiomatic systems feature the rule of *modus ponens*, which requires only the verification of axioms and a single rule of inference. A soundness theorem for a deductive system expresses that all provable sentences are true: All sentences are provable.

Properties of soundness can be either weak or strong, the former of which is a limited version of the other. Strong soundness entails that any sentence P of the language upon which the deductive system is based that is derived from a set (S) of sentences in that language is also a logical consequence of that set. In other words, any system that makes all members of S true will also make P true. Where S is empty, this is a statement of weak soundness.

In arithmetic, T is arithmetically sound if all theorems of T are actually true about the mathematical integers.

The completeness property is the complement of soundness. The completeness property requires that every validity or truth is provable.

Soundness: If something is provable, it is valid. If $\vdash \phi$ then $\models \phi$.

Completeness: If something is valid, it is provable.

If $\models \phi$ then $\vdash \phi$.

Completeness of first-order logic was established by Austrian-born American mathematician and philosopher Kurt Gödel (1906–1978); however, earlier work by the Norwegian mathematician Thoralf Skolem included some of the earliest results in this area.

Gödel suggested two incompleteness theorems of mathematical logic. These works were important to both mathematical logic and the philosophy of mathematics. They are widely interpreted as disproving Hilbert's system for finding a complete and consistent set of axioms for all mathematics as impossible. According to Gödel, you cannot have a set of axioms that are as follows:

1. Complete, that is, for any ϕ , either $A \models \phi$ or $A \models \neg \phi$.
2. Sound, that is, if $A \models \phi$ then ϕ is true of the natural numbers. (It is not possible to prove, e.g., that every number is prime or that $1 = 2$.)
3. Axiomatizable (i.e., that you can write a program that examines a formula and tells you whether or not it is an axiom).

For every assumption you make about natural numbers, there will be elements that are true but cannot be proven.

Gödel's incompleteness theorems were followed by closely related theorems on the limitations of formal systems, including Alfred Tarski's undefinability theorem, Alonzo Church's proof that David Hilbert's *Entscheidungsproblem* is unsolvable, and Alan Turing's theorem that no algorithm can solve the halting problem.

Together, soundness and completeness imply that all and only validities are provable.

Thora L. Fitzpatrick

See also Deduction; Induction; Logic, Formal and Informal; Logic, Inductive; Logical Theory.

Further Readings

- Copi, I. (1979). *Symbolic logic* (5th ed.). Macmillan.
 Hinman, P. (2005). *Fundamentals of mathematical logic*.
 A. K. Peters.
 Smullyan, R. M. (1992). *Gödel's incompleteness theorems*. Oxford University Press.
 Williamson, J. (2010). *Key terms in logic*. Continuum Publishing.

SPECULATION

Viewed purely as a mental operation, *speculation* is performed by way of pragmatic, future-oriented reflection and conjecture which might be based in experience or other empirical evidence but doesn't necessarily have to be. In its classic philosophical sense, *speculation* is concerned with such metaphysical questions as first causes of the universe, whether the soul is immortal; it is a search for ultimate meanings and truths. In popular culture, *speculative* inquiries into such realms as *cryptozoology* and *ancient astronaut hypotheses*, considered pseudo-scientific due to the absence of verifiable data, are often found to be interesting but not credible. (It can be noted that debunkers of theories may themselves be engaging in speculation.)

This entry discusses speculation historically and identifies major theorists, theories and situations related to economic speculation. Game theory, decision theory, and crowdsourcing are described as applications of increasingly sophisticated speculative techniques, with a brief focus on ethical considerations related to certain forms of speculation.

Models of Speculative Reasoning

Despite the heavy emphasis Western cultures place upon logical operations, a high percentage of human decisions are carried out based upon intuition, emotion, experiential knowledge—all of which factor into processes of speculation. Procedures to simplify decision-making in the face of

risk and uncertainty have been developing since the Dark Ages.

When there are multiple options, uncertainty of outcomes produces affective disturbance in proportion to the level of emotional investment in a hoped-for outcome; the more choices, the more emotionally loaded the decision and sense of risk becomes. Difficult outcomes for certain situations that tend to recur in human experience include the *Hobson's choice* (one option, no alternatives: "Take the horse nearest the stable door or walk"); *Morton's fork paradox* (two options which seem opposed have the same, usually unpleasant, end: "If she is a witch, she will float and face execution; if she isn't, she will drown"); *Sophie's choice* (several options, all having undesirable outcomes); *Procrustean bed* (where all options are tied to a single set of arbitrary data which may be violently enforced); and *double-bind situations* (discursively constructed dilemmas framed to preclude any workable response, such as "Are you still lying?").

Attempts to reduce risk and uncertainty in decision-making produced early strategies for more accurate speculation. The Franciscan friar William of Ockham, a 14th-century logician, prescribed that one should apply empirical evidence to theory. If several similar theories are offered, one should select the simplest, that is, one that makes the fewest assumptions. Thus, *Ockham's razor* will shave away extraneous material, leaving the "elegant solution."

The relationship between gambling and risk management is obvious; however, the ability to mathematically speculate upon outcomes was first documented in 1550 when Italian mathematician-physician Girolamo Cardano attempted to calculate the odds of a pair of dice rolling a double-six. In the 17th century, French mathematician-philosopher Blaise Pascal and number theorist Pierre de Fermat also addressed the "problem of points" first posed by the Franciscan monk Luca Pacioli in the late 15th century. The question centers on a hypothetical game of chance which is interrupted with stakes still on the table; the problem is how to divide the pot fairly. Pascal's and Fermat's calculations were foundational for *probability theory*.

In terms of abstract, metaphysical speculation, Pascal's logical reflection focused on the question

of whether God exists. Since the existence of God cannot be determined by reasoning alone, Pascal weighed the benefits accruing from being right against the consequences of being wrong. *Pascal's wager* led him to “decide” to believe in God.

Development of Mathematical Models

The realization that outcomes of apparently random or chance events might mathematically be forecast led to development of increasingly sophisticated mathematical models oriented toward predictive efficacy. The history of the commercial insurance industry, for example, emerged from efforts to manage risk and uncertainty associated with maritime enterprises. Galileo's development of the telescope in 1609 enhanced not only metaphysical speculation about the cosmos but also economic speculation about when (and whether) merchant ships would arrive in port. In the City of London, dissemination of shipping news in Edward Lloyd's coffee house in 1638 evolved into the global insurance corporation Lloyd's of London.

Inspired by Newton's *Principia*, Abraham de Moivre (1677–1754) published *The Doctrine of Chance* in 1718, wherein he applied principles of statistics and probability to games of chance such as “the gambler's ruin.” His ideas of *binomial distribution* provided grounding for the great mathematician Carl Friedrich Gauss's (1777–1855) theorem of *Gaussian distribution, function, and error curve* (the *bell curve*). De Moivre's study of mortality statistics provided a foundation for the *theory of annuities* and actuarial tables used in the life insurance industry. Pierre-Simon Marquis de Laplace (1749–1827) extended probability theory into seminal models of statistics. The deMoivre–Laplace *central limit theorem* was a key advance in probability theory; it posits that the sum or average of a large number of independent but identically distributed variables will approximate a “normal” distribution.

British statistician Thomas Bayes's (1701–1761) major contribution to probability studies is his inductive approach to *probability inference*, based upon outcomes of prior trials. His *Essay Towards Solving a Problem in the Doctrine of Chances* (1764) presented conclusions that were both highly influential and highly debatable. Bayes's

Theorem is a model of *conditional probability* which factors in the interplay of subjective degree of belief and available evidence.

The polymath Sir Francis Galton (1822–1911) used statistics and probability models in his studies of intelligence and developments in psychometrics that included definitions of “the normal curve” and *regression toward the mean* when extreme random events are remeasured. First germane to his study of genetics, the regression concept has been applied (and also misapplied) in economic analysis and stock forecasting, as well as in experimental design, weather forecasting, and prediction of batting averages.

Stochastic processes are those that are “subject to chance” rather than deterministic, as with trials occurring within a framework where initial conditions and, therefore, outcomes are the same. Stochastic models and techniques have been applied to learning and manufacturing processes, biological phenomena, population questions, and even in musical composition. In 1900, French doctoral student Louis Bachelier (1870–1946) successfully defended a unique dissertation at the Sorbonne. His thesis developed a mathematical approach to financial speculation with a seminal application of stochastic analysis. His *theory of speculation* incorporated the notion of *Brownian motion* that would be more fully elucidated by Albert Einstein in a groundbreaking paper published 5 years later. Bachelier's work laid the groundwork for stochastic calculus; the date of his dissertation defense, March 29, 1900, has been called *the birthdate of mathematical finance*.

Innovations in speculation science near the turn of the century incorporated the scholarship of economists Irving Fisher (1867–1947), whose 1907 enunciation of a formula for *net present value* (one discounts cash flow at a rate congruent with the level of risk involved) is a boon to valuation of assets, as well as the theories of Frank Knight (1885–1972). An influential founding member of the Chicago school of economics, Knight was the first to attempt to explain the distinction between *risk* (where outcomes are unknown but measurable, being governed by probabilities) and *uncertainty* (which involves unmeasurable unpredictability). Broadly similar to this concept of *Knightian uncertainty* Nicholas Nassim Kaleb's (1960–) *black swan theory*

discusses the impact of unexpected, improbable events that seem, in hindsight, to have been inevitable.

Fischer Black's (1938–1995) and Myron Scholes's (1941–) work in the growing field of financial economics exerted a strong impact on risk management. The *Black–Scholes pricing model*, published in 1973, is very close to the *Bachelier model*. Robert Merton (1910–2003) did similar work in evaluation of stock options. Scholes and Merton were awarded the Nobel Prize in Economics in 1997. (Because the prize is not presented posthumously, Fischer Black was not eligible.)

The interdisciplinary scholar Roy Radner (1927–) said in 1979, “When traders come to a market with different information about the items to be traded, the resulting market prices may reveal to some traders something about the information available to other traders” (1979, p. 655). This initial information taken collectively will establish an equilibrium of expectations that can ultimately affect prices and trading, and this is the basis of the *rational expectations equilibrium model*. Similarly to the *observer effect* correlated with *Heisenberg's uncertainty principle* in quantum physics, information inputs and expectations of actors can decrease potential exactitude, increase indeterminacy, and must be factored into calculations.

Development of Nonrational Models

The aim of speculation is to manage uncertainty about possible outcomes by weighing probabilities. In 1738, Daniel Bernoulli's reflection on random events based upon the individual's emotional orientation to a proposed outcome laid the groundwork for *risk science*.

Centuries later, cognitive psychologists Amos Tversky and Daniel Kahneman (1974) elaborated on the affective dimension of decision-making—or gut feelings—when they investigated how people actually make decisions in risk situations. They found that the use of three particular heuristic sets to make decisions under uncertainty conditions are predictable. These are *representativeness*, *availability of previous instances*, and *anchoring* (overreliance upon the conceptual starting point)

and *adjustment*. Because of the irrational nature of such an approach, these mental shortcuts will also lead to predictable cognitive biases. Their research demonstrating the tendency of human judgments to follow intuitive, nonrational processes in decision-making rather than to conform to mathematical models was foundational in the field of *behavioral economics*.

Tversky and Kahneman's *prospect theory* delineates the effect that the framing of choices and potential outcomes conceived of terms of “possible gains” as opposed to “losses” will have on the person's decision, with the “loss” frame being weighted more heavily. The theory, developed over three decades, illustrated the difference between subjective perception of probabilities and mathematical probabilities which leads to illogical choices in the face of risk and uncertainty.

Economic Speculation: Bubbles and Crashes

Banks were established during the Roman Empire, but banks could not manage the totality of financial transactions, resulting in the affiliated trade of moneylender. The first stock-traders were itinerant European moneylenders from about the late Middle Ages. A stock exchange existed in Antwerp, Belgium, as early as 1531. The East India companies began issuing shares in the early 1600s in an effort to mitigate the risk and high losses of maritime trade. The London stock exchange began in coffee houses in 1698 (as noted earlier concerning the founding of Lloyd's of London) and is one of the oldest. The New York stock exchange, the world's largest, was established in 1792. Edward A. Calahan's invention of the stock-ticker in 1867 revolutionized the buying and selling of stocks and stimulated stock speculation.

If one arranges arbitrage (buying a commodity at a certain price and immediately re-selling it at a higher price), investment, and gambling along a spectrum, speculation will usually fall somewhere between investment and gambling, although that is not a bright line. Philip Carret in *The Art of Speculation* (1927/2007) described the difference between a “speculator” and an “investor” in this way:

The man who bought United States Steel at 60 in 1915 in anticipation of selling at a profit is a speculator . . . On the other hand, the gentleman who bought American Telephone at 95 in 1921 to enjoy the dividend return of better than 8% is an investor. (p. 5)

Stock-trading is an intrinsically speculative activity, with speculation often leading to overspeculation and booms, bubbles, and crashes.

The stock market crash of 1929 was attributed to overspeculating “on margin” (simply put, financing primarily through credit). In 1936, economist John Maynard Keynes, himself a successful stock-speculator, wrote: “Speculators may do no harm as bubbles on a steady stream of enterprise. But the situation is serious when enterprise becomes the bubble on a whirlpool of speculation.” The “2008 bubble” which burst globally was also blamed on exuberant overexpansion of speculation in oil, food prices, real estate, and other commodities.

Irrational investor speculation led to the *Dutch tulip craze* in 1637, one of the first economic bubbles on record. An account in the April 1876 issue of *Harper's New Monthly Magazine* described the Dutch people's hysteria for the flower having grown so extreme that a single bulb sold for “\$1840, a new carriage, two gray horses, and a complete suit of harness.” Regarding the *bitcoin craze* of 2013, Nout Wellink, former president of the Dutch Central Bank, condemns this as pure speculation, comparing it to the 17th century tulip craze, says, “At least then you got a tulip [at the end], now you get nothing” (quoted in *The Hill*, n.p.).

Game Theory and Decision Theory

In 1944, John von Neumann and Oskar Morgenstern established a mathematical approach to decision-making which assumes that choices will be made by rational, consistent actors. Herbert Simon debunked this in presenting his theory of *bounded rationality*. Under this model, the absence of resources required for fully rational (*optimizing*) decisions requires that intuition, experience, and logic work together to arrive at “good enough” (*satisficing*) decisions. Even so, several

models of decision-making hypothesize a certain level of rationality and have found some application in college and university curricula for modeling and teaching critical thinking.

Speculation framed as rational and strategic is the basis of *game theory*. The three ingredients for a game generally are players, actions or strategies each can take, and each player's preferences. The objective is to logically and objectively determine choices that a player should make to obtain the best possible outcome. Game scenarios can be developed across a variety of human situations, including disputes and negotiations. The scenarios are structured in such a way to maximize the interdependence and interaction of the players, who may perceive that their goals are incompatible, literally the recipe for *conflict*. Game outcomes may be *zero sum*, *positive sum*, or *negative sum*.

In a zero-sum game such as Matching Pennies, each player's gain equals others' losses, so there is no cumulative net gain or loss, and if a party benefits, it's at the expense of the others. In a positive-sum game, the cumulative total is greater than zero, indicating that all parties have benefited in some way in a win-win scenario. A negative-sum game culminates in a lose-lose outcome where all players suffer loss, with the cumulative total being a negative number.

Games can be cooperative or noncooperative. But even in noncooperative games, there is a point of equilibrium where none of the players would benefit if they diverged from an agreed-upon course of action. John Forbes Nash (1928–2015), whose life was the basis for the film *A Beautiful Mind*, was awarded the Nobel Prize in Economics for developing this insight into what has become known as the *Nash equilibrium theory*.

Examples of games built around hypothetical situations include the Prisoner's Dilemma, the Bubble Game, and the Kobayashi Maru game based on the *Star Trek* series. Real-world situations categorized as negative-sum games include war and genocide, arms races, bee colony collapse disorder, and similar environmental situations. Positive-sum games that occur in life could include surplus trade agreements, microloans, corporate joint ventures, and integrative bargaining within organizations.

American ecologist Garrett Hardin (1915–2003) said, “Each man is locked into a system that compels him to increase his herd without limit—in a world that is limited” (1968, p. 1244). Hardin’s *tragedy of the commons* thesis states that individuals will act rationally in their own behalf to maximize personal gain, but this is irrational in terms of long-term outcomes and the common good when they damage or deplete a common resource. A number of game theory scenarios are predicated upon this thesis. One true instance is speculation involving wildlife commodities such as ivory, black rhino horn, and similar animal parts; as the creature moves closer to extinction, the body part increases in value.

Decision theory or *decision science* or *rational choice theory* is related to game theory and is concerned with the ways that people, in a nonrandom way, freely choose among options available to them. Decision science is rooted in scholarship and theory but is generally aimed more toward pragmatic application in economics, statistics, psychology, the social sciences, and especially marketing and strategic communication. Decision theories tend to delineate how people actually do arrive at the decisions they make rather than the underlying processes that, ideally, they should follow.

Speculation as Forecasting

A relationship between crowd psychology and speculation and the impact of suggestion and “herd mentality” on economic and sociopolitical situations has been well documented. But rather than solely negative outcomes whereby crowds can act violently and do damage, recent experiments have demonstrated that crowds possess uncommon wisdom as well. This is beneficial in institutional decision-making, instrumental in public policy formulation, and directive for market research. Combining aspects of game theory and decision science, *social forecasting* or *crowdsourcing* draws on “the wisdom of crowds” and *the law of large numbers*; answers are expressed numerically with calculations facilitated by computer applications.

One such forecasting endeavor is The Good Judgment Project, now in its fourth year of tournaments that ask nonexpert participants to predict

outcomes of world events. University of Pennsylvania psychologist Phil Tetlock (1954–), cofounder of the project with Barbara Mellers and Don Moore, had studied political forecasters for two decades and became interested in how they achieved success or failure. Tetlock identified two categories of predictors. “Hedgehogs” are invested in and very knowledgeable about a “single central vision”—they have bought into a “grand theory” and are “true believers.” “Foxes,” on the other hand, have many interests, are humble about their level of knowledge, and intellectually adaptable. Foxes are better at prediction. Findings from the first three tournaments indicate that ordinary citizens, with a little training in probability and prediction, consistently outperform such professionals as the members of the U.S. Central Intelligence Agency.

Data mining, predictive analytics, and social forecasting are used to understand alternatives, recognize and resolve cognitive biases, and bolster non-data-based decision-making practices. Harvard’s interdisciplinary Decision Science Laboratory serves a pool of investigators, assisting in experiment design and coding, recruiting of subjects, analysis of findings, and training of student-investigators in studies focusing on nonrational elements of decision-making. Similarly, Columbia Business School’s Center for Decision Sciences facilitates research and understanding on consumer behavior, the implications of decision making on public policy, and the neurological underpinnings of judgment and decision-making.

Leslie J. Reynard

See also Evidence; Game Theory; Generalization; Prediction; Theory Construction

Further Readings

- Bachelier, L. (2007). *Louis Bachelier’s theory of speculation: Origins of modern finance* (M. Davis & A. Etheridge, Eds. & Trans.). Princeton University Press (Original work published 1900).
- Carret, P. (2007). *The art of speculation*. Marketplace Books (Original work published 1927).
- Dickenson, M. (2013). Prediction and good judgment: Can ICEWS inform forecasts? *Predictive Heuristics*. Duke University, Ward Labs. Retrieved July 22, 2014, <http://predictiveheuristics.com/2013/11/12/>

- prediction-and-good-judgment-can-icews-inform-forecasts/
- Hagstrom, R., CFA. (2013, February 2013). What is the difference between investing and speculating? *Inside Investing*. CFA Institute. Retrieved June 14, 2014, <http://blogs.cfainstitute.org/insideinvesting/2013/02/27/what-is-the-difference-between-investing-and-speculation-2/>
- Hardin, G. (1968). The tragedy of the commons. *Science*, 162(3859), 1243–1248.
- Mason, C. F., Bulte, E. H., & Horan, R. D. (2012). Banking on extinction: Endangered species and speculation. *Oxford Review of Economic Policy*, 28(1), 180–192.
- Moinas, S., & Pouget, S. (2013). The bubble game: An experimental study of speculation. *Econometrica*, 81, 1507–1539. <http://dx.doi.org/10.3982/ECTA9433>
- Radner, R. (1979). Rational expectation equilibrium: Generic existence and the information revealed by prices. *Econometrica*, 47(3), 655–678. <https://www.jstor.org/stable/1910413>
- Stäehli, U. (2013). *Spectacular speculation: Thrills, the economy, and popular discourse* (E. Savoth, Trans.). Stanford University Press.
- Tversky, A., & Daniel, K. (1974, September 27). Judgment under uncertainty: Heuristics and biases. *Science*, 185(4157), 1124–1131.
- Wellink, N. Quoted in Pry, P. Bitcoins, tulipmania and electric grid insecurity. *The Hill* (March 5, 2021). <https://thehill.com/opinion/national-security/541385-bitcoins-tulipmania-and-electric-grid-insecurity>

STATISTICAL INFERENCE, BAYESIAN

Bayesian statistical inference is based on the subjective interpretation of probability. Uncertainty about an unknown quantity or a parameter of interest is expressed by a probability distribution that represents the agent's subjective degrees of belief. Bayesian Conditionalization and Bayes's Theorem describe how learning data changes this (prior) probability distribution into a posterior probability distribution. Standard procedures of statistical inference, such as point and interval estimation, can be formulated as functions of the posterior distribution. Hypothesis tests are

evaluated by means of the Bayes factor—that is, the degree to which the odds for either hypothesis have changed in the light of the data. All in all, Bayesian statistical inference has high flexibility and can be applied to a large number of important statistical problems, such as generalized linear regression, model selection, and sequential analysis.

Basic Principles

It is well known that in classical, frequentist inference we cannot assign a probability that an unknown parameter μ takes a particular value or falls inside a particular interval. We can only talk about the probability of observing certain types of data, given a particular value of that parameter, and interpret this probability as a relative frequency. Bayesians break with this rule: They use (subjective) probabilities for expressing all kinds of statistical uncertainties, including the probability of an unknown parameter being in a particular interval.

Suppose we test a null hypothesis $H_0: \mu \leq 0$ against an alternative hypothesis $H_1: \mu > 0$. A Bayesian assigns a prior probability distribution over μ (e.g., in the form of a probability density function). Suppose we observe some evidence E , for example, values of an observable whose probability distribution depends on the value of μ . The Bayesian calculates the posterior probability of H_0 and H_1 as the conditional probability of these hypotheses, given the observed evidence: $p_{\text{new}}(H_0) = p(H_0|E)$ and $p_{\text{new}}(H_1) = p(H_1|E)$. This updating principle is called *Bayesian Conditionalization*.

For calculating these conditional probabilities, the Bayesian statistician makes use of *Bayes's Theorem*:

$$p(H_0 | E) = p(H_0) \frac{p(E | H_0)}{p(E)}$$

$$p(H_1 | E) = p(H_1) \frac{p(E | H_1)}{p(E)}$$

It has been argued that the posterior distribution of an unknown parameter is the real target of

statistical inference, and Bayesian inference is the only school of statistics that can deliver it. Moreover, the posterior distribution can be used in a decision-theoretic calculus for inference and decision-making, in agreement with *expected utility theory*. Since the likelihoods $p(E|H_0)$ and $p(E|H_1)$ are based on the sampling distribution of the observed data under the competing hypotheses, the posterior distribution is the product of a distinctly subjective element—the prior distribution—and a more objective element, represented by the likelihoods.

Bayesian statistical inference has some important advantages over frequentist inference. Most importantly, the posterior takes the base rate of an event or hypothesis into account (i.e., $p(H_0)$ and $p(H_1)$). Therefore, it avoids common fallacies in frequentist statistics, such as base rate neglect or the inverse probability fallacy (i.e., inferring from $p(E|H)$ to $p(H|E)$). Bayesians also dispense with p values and do not need to bother about their correct interpretation.

Similar to frequentist tests, Bayesian statistical methods produce point values of parameter estimates (e.g., regression coefficients). Uncertainty around these estimates is quantified in terms of a *credible interval*. Point estimate and credible interval are the central tendency (mean, median, or mode) of the posterior distribution and its 95 central percentiles or 95% highest density interval, respectively. Unlike frequentist confidence intervals, credible intervals have a natural and intuitive interpretation: a 95% credible interval of parameter μ is an interval I in the parameter space such that $p(\mu \in I) = 95\%$. Credible intervals are also at the basis of modern Bayesian inference techniques, such as regions of highest posterior density.

To wrap up, the main ideas of Bayesian inference can be summarized by the following three principles:

- the representation of subjective degrees of belief in terms of probabilities;
- the use of Bayesian Conditionalization and Bayes's Theorem for updating subjective degrees of belief;
- the use of the posterior probability distribution for estimating parameter values, accepting hypotheses and making decisions.

Hypothesis Tests and Bayes Factors

The prior and posterior distributions give clear answers to the questions of what we should *believe* and what we should *do* or *decide*. However, they do not quantify the amount of statistical *evidence* for either hypothesis, as for example p values do in the context of frequentist inference. That is, they do not state whether the evidence favors H_1 over H_0 or vice versa.

Within the Bayesian framework, hypothesis tests are generally based on *Bayes factors*. They are defined as the ratio of posterior and prior odds and describe by how much the evidence E shifts the posterior distribution from the null hypothesis H_0 to the alternative H_1 :

$$B_{10}(E) = \frac{p(H_1|E)}{p(H_0|E)} / \frac{p(H_1)}{p(H_0)}$$

Alternatively, one can calculate Bayes factors as the likelihood ratio between the two competing hypotheses:

$$B_{10}(E) := \frac{p(E|H_1)}{p(E|H_0)}$$

In this reading, the Bayes factor expresses how much more expected the evidence is under H_1 than under H_0 —or in other words, how much it discriminates H_1 from H_0 . For instance, in the case of a t test or regression, there is some evidence in favor of (against) group difference or the inclusion

Table 1 Interpretation of Bayes Factors as Statistical Evidence for a Hypothesis

Bayes Factor BF_{10}	Interpretation
>100	Extreme evidence for H_1
30–100	Very strong evidence for H_1
10–30	Strong evidence for H_1
3–10	Moderate evidence for H_1
1–3	Anecdotal evidence for H_1
1	No evidence for either hypothesis
1/3–1	Anecdotal evidence for H_0
1/3–1/10	Moderate evidence for H_0
1/10–1/30	Strong evidence for H_0
1/30–1/100	Very strong evidence for H_0
<1/100	Extreme evidence for H_0

of a coefficient in a model when the probability of the observed data is higher (lower) for the model that allows group difference or includes the coefficient with respect to the model with equal groups or without the coefficient; $BF_{10} > 1$ ($BF_{10} < 1$). A conventional interpretation of Bayes factors is provided in Table 1.

An important feature of Bayes factors is that they can express evidence for the null hypothesis—something that is impossible in frequentist inference where evidence either speaks against the null hypothesis or is inconclusive. By being able to quantify the degree of support for the null hypothesis, Bayesian statistics makes a contribution to fighting the file drawer effect: That is, *nonsignificant results* have a hard time to be published in scientific journals. The impossibility of distinguishing between inconclusive results and results that support the null hypothesis in frequentist statistics contributes to this effect since evidence in favor of the null is systematically suppressed. Bayesian statistics, by contrast, allows to confirm theoretically important null hypotheses. Therefore, it is less susceptible to the file drawer effect and contributes indirectly to combatting the replication crisis.

Challenges

Bayesian statistics are often called “too subjective for science.” The idea is that the use of subjective degrees of belief in inference is at odds with the neutrality and impartiality of scientific reasoning, where one should “let the data speak for themselves.” Think, for example, of a regulatory agency that needs to decide on the admission of a drug or vaccine against COVID-19. Certainly it would not be impressed by a probability distribution that is mostly grounded on the subjective degrees of belief of the experimenters (who usually have a conflict of interest with respect to the study results). This is often called the *problem of subjectivity* or the *problem of the prior*.

Priors used in statistical analyses can roughly be categorized in three groups. Subjective priors embody the subjective beliefs and/or prior knowledge of the researcher, for instance a normal distribution that represents the expected value of a certain parameter and the uncertainty around it. Default priors or objective priors are distributions

that satisfy certain general desiderata instead of particular beliefs or prior knowledge the researcher might have. Empirical priors are distributions based on a (certain number of) small subsample(s) of the data to be analyzed.

It is the prior, in any form, that lies at the root of the “too subjective for science” criticism. The responses are twofold: Some Bayesians defer to default or objective priors for inference, while others defend the use of properly subjective priors as scientifically sound. After all, the subjectivity inherent in the specification of priors is explicit, allows others to substitute their own priors, and affords robustness analyses across priors instantiating different prior beliefs. In short, unlike many other decisions made by researchers during the course of an experiment, the subjectivity introduced by the prior in Bayesian statistical analyses is open to inspection and evaluation. Finally, it has been pointed out that subjectivity is inherent to any school of statistical inference and that similar objections can also be raised against the frequentist paradigm.

Applications

Bayesian statistics encompasses counterparts to all standard statistical tests (t test, ANOVA, regression, etc.) and advanced statistical modeling methods (mixed models, hierarchical models, structural equation models, etc.). Software packages such as SPSS and Stata offer Bayesian statistical analyses, in addition to classical/frequentist analyses. Furthermore, there is free-to-use software dedicated to Bayesian statistics, such as the programs MCSim, OpenBUGS, and JAGS, the probabilistic programming language Stan, R packages such as LaplacesDemon, and Python packages such as ArviZ and PyMC3. Finally, there is the free and open-source graphic user interface software package JASP that offers both frequentist and Bayesian analyses, both standard and advanced.

Bayesian statistical inference has several additional advantages, especially when conducted with Bayes factors:

Automatic parsimony. As the prior probability of a hypothesis is distributed across parameters and parameter values (i.e., prior distribution), the more complex a model is (i.e., more parameters and range of values), the more the prior density is

spread out. For the Bayes factor, the likelihood of the hypothesis is marginalized over its parameter, which decreases rapidly when parameters and parameter values are included that do not match the data.

Comparing non-nested hypotheses. Because only the likelihoods of hypotheses are considered for the Bayes factor, no distinction is made between nested and non-nested models. This allows, for instance, the evaluation of order-constrained hypotheses such as $H_1: \mu > 0$ versus $H_2: \mu < 0$.

Sequential analysis. Bayesian hypothesis does not require special corrections for sequential data collection and analysis. The symmetry of the Bayes factor and its sole reliance on likelihoods make the rules of data collection and optional stopping irrelevant to data interpretation. Thus, one can monitor evidence as it accumulates (e.g., for or against H_0) without worrying about the stopping rule. This feature of Bayesian inference is also contentious since according to critics, it licenses reasoning to a foregone conclusion and neglects the (preexperimental) error probabilities of an inference.

Incorporation of existing knowledge. The inclusion of a prior in Bayesian analysis allows the incorporation of existing knowledge, mimicking learning as we know it. Existing knowledge about a phenomenon (e.g., efficacy of a drug) can be captured in a prior and thus contribute to the analysis. Consequently, the resulting posterior can inform the prior distribution for the next experiment. As Dennis Lindley put it, “today’s posterior is tomorrow’s prior.”

Flexible research planning. While the Bayesian framework is not concerned with power analysis in the frequentist sense, it does afford the research planning that power analyses provide. Monte Carlo simulations allow us to draw random samples across scenarios (e.g., where H_0/H_1 is true), priors, sample sizes, and decision boundaries (e.g., $BF_{10} > 10$) to investigate the percentage of samples in which a confirmatory decision is reached for a given statistical test. Conversely, we can determine the necessary sample size for a given statistical test, hypothesis, prior, decision boundary, and probability for reaching this decision boundary (e.g., obtaining a $BF_{10} > 10$ with a probability of .8). Such Bayes factor design

analysis can be used as easily for sequential designs as for fixed sample size designs.

Jan Sprenger and Noah van Dongen

See also Biostatistics; Induction; Inference; Statistical Inference, Frequentist; Statistics

Further Readings

- Bernardo, J. M., & Smith, A. F. M. (1994). *Bayesian theory*. Wiley.
- Edwards, W., Lindman, H., & Leonard J Savage, L. J. (1963). Bayesian statistical inference for psychological research. *Psychological Review*, 70, 193–242.
- Howson, C., & Urbach, P. (2006). *Scientific reasoning: The Bayesian approach* (3rd ed.). Open Court.
- Jeffreys, H. (1961). *Theory of probability* (3rd ed.). Oxford University Press.
- Kruschke, J. K., & Liddell, T. M. (2018). The Bayesian new statistics: Hypothesis testing, estimation, meta-analysis, and power analysis from a Bayesian perspective. *Psychonomic Bulletin & Review*, 25, 178–206.
- Lee, M. D., & Wagenmakers, E.-J. (2013). *Bayesian cognitive modeling: A practical course*. Cambridge University Press.
- Mayo, D. G. (2018). *Statistical inference as severe testing: How to get beyond the science wars*. Cambridge University Press.
- Rouder, J. N., Speckman, P. L., Sun, D., Morey, R. D., & Iverson, G. (2009). Bayesian t tests for accepting and rejecting the null hypothesis. *Psychonomic Bulletin & Review*, 16, 225–237.
- Royall, R. (1997). *Statistical evidence: A likelihood paradigm*. Chapman & Hall.
- Schönbrodt, F. D., & Eric-Jan Wagenmakers, E.-J. (2018). Bayes factor design analysis: Planning for compelling evidence. *Psychonomic Bulletin & Review*, 25, 128–142.
- Sprenger, J. (2018). The objectivity of subjective Bayesianism. *European Journal for Philosophy of Science*, 8, 539–558.
- Sprenger, J., & Hartmann, S. (2019). *Bayesian philosophy of science*. Oxford University Press.

STATISTICAL INFERENCE, FREQUENTIST

Probability theory, a branch of mathematics, provides the foundation and the overarching framework for statistical inference. As such, statistics is

in essence applied mathematics and should not be treated as a collection of recipe-like applications of computer-driven statistical techniques without any real understanding of their assumptions, limitations, proper implementation, and warranted interpretation of their results. This entry aims to summarize the basic results of frequentist inference in precise mathematical notation so as to avoid possible sources of confusion and misapprehensions that can result, according to critics, from texts that omit or water down such formulations.

Model-Based Frequentist Inference

Very few revolutions in scientific thought can be dated with any certainty, but *modern frequentist inference* is an exception. It was single-handedly founded by Ronald A. Fisher (1890–1962), who initiated a change of paradigms in statistics by recasting Karl Pearson’s descriptive statistics and curve fitting, into a model-based statistical induction that revolves around the concept of a prespecified parametric statistical model $M\theta(x)$. Before Fisher, the notion of a statistical model was implicit, and its role was primarily confined to the *description* of the distributional features of *the data in hand* $x_0 := (x_1, \dots, x_n)$, implicitly assumed to be independent and identically distributed (IID), using a frequency curve that belongs to the Pearson family. Unlike Pearson, who would commence with data x_0 seeking the best fitting frequency curve to approximate the histogram of x_0 , Fisher proposed to begin with (a) a prespecified statistical model (a hypothetical infinite population) and (b) view x_0 as a typical realization thereof: “Of what population is this a random sample?” (Fisher, 1922, p. 313), emphasizing that “the adequacy of our choice may be tested a posteriori” (Fisher, 1922, p. 314). A statistical model $M\theta(x)$, comprising the set of probabilistic assumptions imposed on one’s data x_0 , describes a stochastic mechanism that could have given rise to x_0 .

Prespecified parametric statistical model. Fisher’s concept of a prespecified statistical model can be formalized as a particular parameterization ($\theta \in \Theta$) of the observable stochastic process $\{X_k, k \in \mathbb{N} := (1, 2, \dots, n, \dots)\}$ underlying data x_0 , whose generic form is:

$$M\theta x = fx; \theta, \theta \in \Theta, x \in \mathbb{R}^{Xn}, \quad \text{for } \Theta \subset \mathbb{R}^m, m < n, \quad (1)$$

where $f(x; \theta)$ denotes the (joint) *distribution of the sample* $X := (X_1, \dots, X_n)$, $\mathbb{R}^{Xn}$ the sample space, and Θ the parameter space. $M\theta(x)$ plays a pivotal role in frequentist inference since it determines: (a) what constitutes a legitimate event on which $f(x; \theta)$ assigns probabilities, (b) appropriate data x_0 , and (c) legitimate hypotheses, and (d) optimal inference procedures; see Spanos (2019).

Specification. $M\theta(x)$ is selected to provide an idealized description of the stochastic mechanism that gave rise to data x_0 with a view to pose the substantive questions of interest framed in terms of θ . The choice of $M\theta(x)$ is based on selecting the appropriate probabilistic assumptions relating to $\{X_k, k \in \mathbb{N}\}$ to render x_0 a typical realization thereof, or equivalently $M\theta(x)$ accounts for all the chance regularity patterns exhibited by data x_0 ; this is established using misspecification testing. Chance regularities constitute recurring patterns in data x_0 that can be accounted for (modeled) using probabilistic assumptions from three broad categories: distribution, dependence, and heterogeneity; see Spanos (2019).

Example 1. The simple Normal model:

$$M\theta(x): X_t \sim \text{NIID}(\mu, \sigma^2), \theta := (\mu, \sigma^2) \in \mathbb{R} \times \mathbb{R}_+, t \in \mathbb{N}, \quad (2)$$

where $\mathbb{R} := (-\infty, \infty)$, $\mathbb{R}_+ := (0, \infty)$, and “ $\sim \text{NIID}(\mu, \sigma^2)$ ” stands for “distributed as Normal IID, with mean μ and variance σ^2 .”

Example 2. The simple Bernoulli model:

$$X_k \sim \text{BerIID}(\theta, \theta(1 - \theta)), x_k = 0, 1, \quad 0 \leq \theta < 1, k \in \mathbb{N}, \quad (3)$$

where “BerIID” denotes Bernoulli IID.

Example 3. The linear regression (LR) model:

$$Y_t = \beta_0 + \beta_1 x_t + u_t, u_t | X_t = x_t \sim \text{NIID}(0, \sigma^2), \theta := (\beta_0, \beta_1, \sigma^2) \in \mathbb{R}^2 \times \mathbb{R}_+, t \in \mathbb{N} \quad (4)$$

The **frequentist interpretation** of probability underlying statistical inference invokes the simple Bernoulli model in (3) and associates the probability of a binary event $A = \{X = 1\}$, $P(A)$, with the *limit* of relative frequencies $1_n = 1/n \sum_{i=1}^n X_i$ as $n \rightarrow \infty$, by appealing to the Strong Law of Large Numbers.

The *long-run metaphor* usually employed to conceptualize probability in terms of relative frequencies, does not *identify* probability with relative frequency; relative frequencies can only approximate (reflect) probabilities; see Spanos (2013).

Statistical inference. The *main objective* of model-based $[M\theta(x)]$ frequentist inference is to learn from data x_0 by *narrowing down* Θ as much as possible, ideally to a single point $\theta = \theta^*$, where θ^* denotes the *true* value of θ in Θ ; shorthand for saying that $M^*(x) = \{f(x; \theta^*)\}$, $x \in \mathcal{R}X_n$, could have generated data x_0 . In practice, $\theta = \theta^*$ is unlikely to be attained, except by happenstance, but that does not preclude learning from x_0 .

A key concept in frequentist inference that underlies both estimation and testing is the *likelihood function*:

$$L(\theta; x_0) = l(x_0) \cdot f(x_0; \theta), \theta \in \Theta, \quad (5)$$

where $l(x_0) > 0$ denotes a proportionality constant. Fisher defined the *maximum likelihood* (ML) estimator $\theta_{ML}(X)$ of θ to be the one that maximizes $L(\theta; x_0)$. He was also the first to draw a sharp distinction between the *estimator* $\theta(X)$ and the *estimate* $\theta(x_0)$ and emphasized the importance of using the sampling distribution of $\theta(X)$ to evaluate the effectiveness of the inference in terms of optimal properties. Neyman and Pearson (1933) introduced the idea of a likelihood ratio, $[L(x_0; \theta_1)/L(x_0; \theta_0)]$, for the null (θ_0) and alternative (θ_1) value of θ , as a basis for constructing optimal tests.

In principle, the *sampling distribution* $f(y_n; \theta)$, of a statistic (estimator, test, predictor), say $Y_n = g(X_1, X_2, \dots, X_n)$, is derived via:

$$F_n(Y_n \leq y) = \dots \{x: g(x) \leq y\} f(x; \theta) dx, \quad \forall y \in \mathcal{R}, \quad (6)$$

where $f(x; \theta)$ is given by the choice of $M\theta(x)$. In (6), the value of θ is always *prespecified*, taking two different forms relating to the underlying reasoning:

(i) *Factual* (point and interval estimation, prediction): evaluated under $\theta = \theta^*$, whatever value θ^* happens to be in Θ .

(ii) *Hypothetical* (testing): evaluated under various hypothetical scenarios based on θ taking different prespecified values relating to the null and alternative hypotheses, $H_0: \theta \in \Theta_0$ versus $H_1: \theta \in \Theta_1$, where Θ_0 and Θ_1 constitute a partition of Θ ; see Spanos (2019).

Estimation (point). The effectiveness of a point estimator $\theta(X)$ is evaluated in terms of how *accurately* pinpoints θ^* , using its optimal properties relating to its sampling distribution, $f(\theta(x); \theta)$, $x \in \mathcal{R}X_n$.

Example 1. For the simple Normal model in (2):

$$\begin{aligned} \text{(i)} \quad X_n &= 1n_k = 1nX_k \sim N_{\mu}, \sigma^2_n, \\ \text{(ii)} \quad s^2 &= 1(n-1)_k = 1n(X_k - X_n)^2 \sim \\ &\quad \sigma^2_n - 1\chi^2(n-1), \end{aligned} \quad (7)$$

where $N_{\cdot}$, and $\chi^2(\cdot)$ denote the Normal and chi-square distributions. These are *optimal* (effective in pinpointing θ^*) estimators of $\theta = \mu, \sigma^2$ since their sampling distributions in (7) indicate that they are both consistent, unbiased, sufficient, and X_n is also fully efficient; see Cox and Hinkley (1974). What is often not appreciated sufficiently is that the notation in (7) is misleading since the underlying reasoning is invariably omitted. For instance, (i) would be stated (under $\theta = \theta^*$) as:

$$X_n = 1n_i = 1nX_i \sim \theta = \theta^* N(\mu^*, \sigma^{*2}_n). \quad (8)$$

One might object to the unnecessary notation in (8), but then one needs why $E(X_n) = \mu$. The latter stems from the unbiasedness of X_n , which is defined by $E(X_n) = \mu^*$, involving factual reasoning.

Confidence interval (CI). A two-sided CI for θ takes the form of a random interval $L(X), U(X)$, where $L(X)$ is the lower bound and $U(X)$ the upper bound, aiming to overlay (cover) θ^* at some prespecified coverage probability $(1 - \alpha)$; one-sided CIs can also be defined by $L(X), \infty$ and $-\infty, U(X)$. The effectiveness of a CI is evaluated in terms of its expected length $E[U(X) - L(X)]$ for a prespecified coverage probability $(1 - \alpha)$. An optimum CI is the one with the minimum expected length for a given $1 - \alpha$; see Neyman (1937).

Example 1. For (2), the $1 - \alpha$ CI takes the form:

$$\begin{aligned} P(X_n - c\alpha^2(s_n) \leq \mu < X_n + c\alpha^2(s_n); \\ \theta = \theta^*) &= 1 - \alpha, \end{aligned} \quad (9)$$

where $s^2 = 1n - 1i = 1n(X_i - X_n)^2$, stemming from the distribution of the pivot:

$$\tau(X; \mu) = n(X_n - \mu)s \sim \theta = \theta^* St(n-1), \quad (10)$$

where “ St ” denotes the Student’s t distribution. Its expected length is:

$$\begin{aligned} E X_n + c\alpha^2 s_n - X_n + c\alpha^2 s_n = \\ 2c\alpha^2 \sigma_n \rightarrow 0 \text{ as } n \rightarrow \infty. \end{aligned}$$

Prediction is similar to estimation in terms of its underlying factual reasoning, but it differs from it in so far as it aims to find the best value of X_k beyond the observed data, say x_{n+1} .

Example 1. In light of (8), an optimal predictor of X_{n+1} is given by $X_{n+1} = X_n$, whose reliability is evaluated using the sampling distribution of the prediction error:

$$x_{n+1} - (X_{n+1} - X_n) \sim \theta = \theta^* N(0, \sigma^2(1 + 1/n)), \quad (11)$$

to construct a $1 - \alpha$ prediction interval:

$$PX_n - c\alpha 2s(1 + 1/n) \leq X_{n+1} \leq X_n + c\alpha 2s(1 + 1/n); \theta = \theta^* = 1 - \alpha. \quad (12)$$

Neyman–Pearson (N-P) testing poses questions on whether $\theta^* - \theta_0$, where θ_0 is the null value, is *statistically negligible* to data x_0 in terms of H_0 and H_1 by partitioning the parameter space Θ and operationalizing *statistically negligible* using error probabilities; see Neyman and Pearson (1933).

Example 1. For testing the following hypotheses in the context of (2):

$$H_0: \mu \leq \mu_0 \text{ vs } H_1: \mu > \mu_0, \quad (13)$$

an optimal Uniformly Most Powerful (UMP) test is defined by:

$$T\alpha := [\tau X = nX_n - \mu_0 s, C1\alpha = x: \tau x > c\alpha], \quad (14)$$

where $C1(\alpha)$ is the rejection region at significance level α . The N-P accept/reject H_0 rules relate to whether x_0 belongs to the acceptance $C0(\alpha) = \{x: \tau(x) \leq c\alpha\}$ or the rejection region.

The sampling distribution of $\tau(X)$ evaluated under H_0 (hypothetical) is:

$$\tau(X) = n(X_n - \mu_0)s \sim \mu = \mu_0 St(n-1), \quad (15)$$

is used to evaluate the Type I error probability and the p value:

$$P(\tau(X) > c\alpha; \mu = \mu_0) = \alpha, P(\tau(X) > \tau(x_0); \mu = \mu_0) = p(x_0). \quad (16)$$

The sampling distribution of $\tau(X)$ evaluated under H_1 (hypothetical) is:

$$\tau X = nX_n - \mu_0 \sigma \sim \mu = \mu_1 St\delta_1; n-1, \delta_1 = n\mu_1 - \mu_0 \sigma, \text{ for all } \mu_1 > \mu_0, \quad (17)$$

where δ_1 is the noncentrality parameter, is used to evaluate the power of $T\alpha$:

$$P(\mu_1) = P(\tau(X) > c\alpha; \mu = \mu_1), \text{ for all } \mu_1 > \mu_0, \quad (18)$$

which in conjunction with α defines its optimality; see Lehmann and Romano (2005).

The effectiveness of an N-P test $T\alpha$ is evaluated using the sampling distributions (15) and (17) of a test statistic $\tau(X)$, with its optimality calibrated in terms of a prespecified Type I error probability α , in conjunction with its power. For instance, the t test in (14) is UMP; see Lehmann and Romano (2005).

The test statistic $\tau(X)$ in (15) differs from the pivot $\tau(X; \mu)$ in (10) in two respects. First, μ is unknown in $\tau(X; \mu)$, but μ_0 in $\tau(X)$ is a known hypothetical value. Second, the underlying reasoning in (15) is hypothetical, under $\mu = \mu_0$, but in (10) is factual, under $\mu = \mu^*$. Having said that, the family resemblance between (15) and (10) provides the basis for the mathematical duality between optimal CIs and optimal N-P tests. For instance, using a pair of the test statistic and a pivot, a UMP test will give rise to a Uniformly Most Accurate (UMA) CI in the one-sided case, and a UMA Unbiased CI in the two-sided case as in (9).

N-P testing differs from Fisher's significance testing in two respects. The first is that adding an alternative hypothesis H_1 to the null H_0 brings into the testing the whole of Θ in conjunction with the error probabilities (Type I, II, power) that underlie the optimal theory of testing analogous to Fisher's optimal theory of estimation; see Fisher (1925). The second is that Fisher's p value is a post-data (x_0 is known) error probability that provides additional information beyond the accept/reject H_0 results.

Frequentist Error Probabilities

Regrettably, most statistics books use the vertical bar (|) instead of the semicolon (;) for the error probabilities in (16)–(18), misleadingly suggesting that they are *conditional* on H_0 or H_1 ; see Cohen (1994). Given that H_0 and H_1 are framed in terms of $\theta \in \Theta$, and θ is an unknown constant, the alleged conditioning is a misconception; see Spanos (2019). Another serious misconception is that error probabilities are often conflated with

the long-run relative frequencies associated with the metaphor underlying the frequentist interpretation of probability. For a sufficiently large number N of replications $x_1, x_2, \dots, x_N$, generated by the same $M\theta(x)$ as the original data x_0 , one can provide an empirical counterpart to the abstract error probabilities using the relative frequencies associated with the metaphor. It is crucial to emphasize, however, that error probabilities are genuine probabilities that calibrate the effectiveness of inference procedures, and the metaphoric long-run relative frequencies represent a way to envision such probabilities in more tangible terms; see Spanos (2013).

Inferential Claims for Estimation and Testing

What is often insufficiently appreciated in statistical inference are the inferential claims warranted by the different forms of inference as they relate to learning from data x_0 about θ^* . For instance, what claim does an optimal estimator $\theta(X)$ of θ can sanction about how the *estimate* $\theta(x_0)$ relates to θ^* ? Despite the widely held belief among textbook writers touting the use of effect sizes, the claim that $\theta(x_0)$ approximates close θ^* , denoted by $\theta(x_0) \simeq \theta^*$, is unwarranted, since $\theta(x_0)$ represents a single value over a wide range of possible values associated with $f(\theta(x); \theta)$, $x \in RX_n$.

A formal way to account for the uncertainty relating to $\theta(X)$ is to construct a CI for θ using $f(\theta(x); \theta)$, $x \in RX_n$, say:

$$P(L(X) \leq \theta < U(X); \theta = \theta^*) = 1 - \alpha, \quad (19)$$

as in (9). Unfortunately, the coverage probability $1 - \alpha$ in (19) relating to that uncertainty does not extend to the observed CI, $CI(x_0) = [L(x_0) \leq \theta < U(x_0)]$, since the latter either includes or excludes θ^* , post-data. Intuitively, one cannot assign a probability to $CI(x_0)$ since no random variable is involved. That is, the pre-data mathematical duality between CIs and N-P testing does not imply inferential duality, post-data. This renders unwarranted any (direct or indirect) probabilistic or likelihoodist assignment to values of θ within $CI(x_0)$; see Cumming (2012).

It is well known that the N-P testing accept (reject) H_0 does not warrant the claim that there is evidence for H_0 (H_1), and neither does a large

(small) p value. The main reason is that the process of going from an accept/reject H_0 result to post-data evidence for or against H_0 or H_1 warranted by data x_0 and test T_α needs to take into consideration the relevant statistical context:

$$(i) M\theta x, (ii) H_0: \theta \in \Theta_0 \text{ vs } H_1: \theta \in \Theta_1, \\ (iii) \text{ data } x_0 \quad (20)$$

Equation (20) includes the statistical adequacy of $M\theta(x)$, the framing of H_0 and H_1 , the test T_α as well as its power. This renders the commonly followed practice of detaching a testing result from (20) and reporting significance/nonsignificance using asterisks highly dubious. This is because all these results are vulnerable to the fallacies of acceptance and rejection; see Mayo and Spanos (2006).

Post-data severity. To account for (20) and address these fallacies, Mayo and Spanos (2006) propose an evidential interpretation of testing results based on post-data severity evaluation. In contrast to CIs, severity is based on post-data hypothetical reasoning, which provides a most potent tool for interrogating the data x_0 with a view to learn about θ^* since it includes an infinity of scenarios with the relevant error probabilities being equally ascertainable pre-data and post-data.

Reliable Inference and Trustworthy Evidence

What are the key steps to ensure reliable inference and trustworthy evidence?

Step 1. One begins with one or more questions of substantive interest and an appropriate data set x_0 given rise to by the stochastic phenomenon of interest.

Step 2. One studies several data plots to detect the chance regularities exhibited by x_0 (Spanos, 2019), and proceeds to (a) select an appropriate statistical model $M\theta(x)$ comprising probabilistic assumptions from three broad categories, distribution (e.g., N), dependence (e.g., I), and heterogeneity (e.g., ID), that could account for all the regularities in x_0 , as well as (b) adopt a parameterization $\theta \in \Theta$ in terms of which to frame the substantive questions of interest.

Step 3. Test the validity of the model assumptions (e.g., NIID) using misspecification testing. If any of these probabilistic assumptions are *invalid*,

then one should *respecify* $M\theta(x)$ (to account for the departures from the original assumptions) until a statistically adequate model is reached. This step is particularly crucial because Steps 4 through 6 presume the statistical adequacy of $M\theta(x)$, without which all inferences based on $f(\theta ML(x); \theta)$, $x \in RX_n$, are likely to be unreliable; estimators could be inconsistent, and the nominal error probabilities for CIs and optimal tests are likely to be very different from the actual ones; see Spanos (2019).

Step 4. One would proceed to derive the likelihood function, $L(\theta; x_0) \propto f(x_0; \theta)$, $\forall \theta \in \Theta$, which is then used to derive the ML estimator $\theta ML(X) = h(X)$ of each θ .

Step 5. Since $\theta ML(X) = h(X)$ is a function of X , its sampling distribution, $f(\theta ML(x); \theta)$, $x \in RX_n$, can be derived via (6), and that determines its optimal properties, such as unbiasedness, full efficiency, sufficiency, and consistency.

Step 6. Using $\theta ML(X)$, an *optimal* estimator of θ , one can proceed to draw inferences relating to learning from data about θ^* , by constructing CIs, testing hypotheses, and predicting, as described above.

In conclusion, reliable and effective inferences that would give rise to trustworthy evidence depend on two preconditions: (i) $M\theta(x)$ is statistically adequate (its probabilistic assumptions are valid for data x_0) and (ii) n is large enough for the inference procedure to have sufficient effectiveness to shed light on the substantive questions of interest.

Aris Spanos

See also Statistical Inference, Bayesian

Further Readings

- Billingsley, P. (1995). *Probability and measure* (4th ed.). Wiley.
- Binmore, K. (1992). *Fun and games; a text on game theory*. Heath.
- Cohen, J. (1994). The earth is round ($p < .05$). *American Psychologist*, 49, 997–1003.
- Cox, D. R., & Hinkley, D. V. (1974). *Theoretical statistics*. Chapman & Hall.
- Cumming, G. (2012). *Understanding the new statistics*. Routledge.
- Doob, J. L. (1953) *Stochastic processes*. Wiley.
- Fisher, R. A. (1922). On the mathematical foundations of theoretical statistics. *Philosophical Transactions of the Royal Society A*, 222, 309–368.

- Fisher, R. A. (1925). Theory of statistical estimation. *Proceedings of the Cambridge Philosophical Society*, 22, 700–725.
- Lehmann, E. L., & Romano, J. P. (2005). *Testing statistical hypotheses* (3rd ed.). Wiley.
- Mayo, D. G., & Spanos, A. (2006). Severe testing as a basic concept in a Neyman–Pearson philosophy of induction. *British Journal for the Philosophy of Science*, 57, 323–357.
- Neyman, J. (1937). Outline of a theory of statistical estimation based on the classical theory of probability. *Philosophical Transactions of the Royal Statistical Society of London, A*, 236, 333–380.
- Neyman, J., & Pearson, E. S. (1933). On the problem of the most efficient tests of statistical hypotheses. *Philosophical Transactions of the Royal Society, A*, 231, 289–337.
- Spanos, A. (2013). A frequentist interpretation of probability for model-based inductive inference. *Synthese*, 190, 1555–1585.
- Spanos, A. (2019). *Probability theory and statistical inference: Empirical modeling with observational data* (2nd ed.). Cambridge University Press.
- Yule, G. U. (1916). *An introduction to the theory of statistics* (3rd ed.). Griffin.

STATISTICS

Our lives abound with data, from the grades of standardized tests to the indices on medical reports, from the number of friends on Facebook to the voting rates in presidential elections. A set of data is not very useful in itself; however, it needs to be analyzed and interpreted through data analysis, the process of examining what data mean and transforming data into useful information. This entry aims to introduce the science called *statistics*. Statistics, the science of collecting, organizing, describing, analyzing, and interpreting data, is an applied branch of mathematics associated with data-analytic procedures. Therefore, statistical data analysis is referred to as the process of examining what data mean, with the goal of using imperfect information—which is the collected data—to understand facts, make decisions, and make predictions about the world. It has applications to virtually every area in our lives. **This entry provides an overview of basic statistics as well as a foundation for understanding other**

entries on statistical inference in this encyclopedia.

Statistics used to be referred to as *any numerical indicator of a set of data* and was equated with “descriptive statistics” until Michael Cowles (1989) pointed out that statistics as a science is not only the practice of collecting and collating quantitative data but also the process of making inferences and drawing conclusions from them. Thus, making inferences is part of how numerical information is processed. Therefore, two elements constitute this section: *descriptive statistics*, which is about summarizing data using numbers or plots, and *inferential statistics*, which concerns using data to make conclusions or decisions. Whereas statistics is a science, a statistic is a number or value measured within some particular context. In order to carry out sensible and meaningful statistical procedures, it is important to understand (a) the context, (b) how and why these data were collected, (c) on which individuals or entities the data were collected, and (d) what questions to answer from the data.

Article I. Data and Data Sets

First, we need to know what data and data sets are; let us define some key terms. An *observational unit* is the person or thing on which measurements are taken, for which a number of different terms are commonly used. A *variable* is a measured characteristic of an observational unit. An *observed value or observation* is the actual value of a variable for every observational unit in a data set. Suppose we have a data set which are the SAT scores of 100 high school students. In this data set, the observational unit consists of the high school students, the variable is the SAT score, and the observed values are the 100 specific scores obtained by the students. In practice, data sets are often arranged in matrices, with rows for each observational unit and columns for variables.

Article II. Descriptive Statistics

Section 2.01 Measures of Central Tendency

One important way to describe statistics is by the measure of *central tendency*, also referred to as the *representative values*, which describe the center point of data distribution. Generally speaking,

data can be divided into two types: *categorical (or nominal) data* and *continuous (quantitative) data*. The distinction between these two types of data has implications for describing and making inferences from the data. Continuous data, also referred to as quantitative data, are data where the values can change continuously, and every little change in value provides information. Examples of continuous variables are height, weight, price, and rate. Continuous data can be rank ordered by value. In contrast, categorical data are used to make distinctions between different groups. Examples of categorical variables are type, gender, and ethnicity. Ranking categorical variables does not make conceptual sense, as no category is superior or inferior to another.

Data by itself can be noisy. For instance, Table 1 shows the prevalence of self-report obesity data among U.S. adults collected by the Center of Disease Control and Prevention (CDC, 2014). Here, listed are the obesity rate of 50 states plus the District of Columbia, but limited information can be obtained just from these data themselves. *Descriptive statistics* is used to construct simple descriptions about the characteristics of a set of quantitative data. An efficient way to describe the data is using *numerical indicators* or *summary statistics*, which are condensed data that best summarizes the data set. There are five important indicators to summarize data: minimum, first quartile, median, third quartile, and maximum. In this case, let us rearrange the data in increasing order, from the state of Colorado with the lowest obesity rate (20.5%), which is referred to as the *minimum (Min)*, to the state of Louisiana with the highest obesity rate (34.7%), which is referred to as the *maximum (Max)*. The state located at the midpoint between Colorado and Louisiana is Maryland, which ranks 26th among these 51 states and has a 27.60% obesity rate, the *median (Md or Mdn)*. Since median divides a distribution of quantitative data exactly in half, it is referred to as a positional or order measure. We can also summarize the data using the *first quartile*, a point with a quarter of values lower than it, and *third quartile*, a point with three quarters of values smaller than it. In this example, the first quartile is Connecticut, ranking the 13th (25.6), and the third quartile is Ohio, ranking the 39th (30.1%). Suppose the data collected are only for 49 states

Table 1 Prevalence of Self-Report Obesity among U.S. Adults (CDC, 2014)

<i>State</i>	<i>Prevalence</i>	<i>95% Confidence Interval</i>	<i>State</i>	<i>Prevalence</i>	<i>95% Confidence Interval</i>
Alabama	33.0	(31.5, 34.4)	Missouri	29.6	(28.0, 31.2)
Alaska	25.7	(23.9, 27.5)	Montana	24.3	(23.1, 25.5)
Arizona	26.0	(24.3, 27.8)	Nebraska	28.6	(27.7, 29.6)
Arkansas	34.5	(32.7, 36.4)	Nevada	26.2	(24.3, 28.1)
California	25.0	(23.9, 26.0)	New Hampshire	27.3	(25.8, 28.8)
Colorado	20.5	(19.5, 21.4)	New Jersey	24.6	(23.6, 25.6)
Connecticut	25.6	(24.3, 26.9)	New Mexico	27.1	(25.9, 28.3)
Delaware	26.9	(25.2, 28.6)	New York	23.6	(22.0, 25.1)
District of Columbia	21.9	(19.8, 24.0)	North Carolina	29.6	(28.5, 30.7)
Florida	25.2	(23.6, 26.7)	North Dakota	29.7	(27.9, 31.4)
Georgia	29.1	(27.4, 30.8)	Ohio	30.1	(29.0, 31.2)
Hawaii	23.6	(22.0, 25.1)	Oklahoma	32.2	(30.8, 33.6)
Idaho	26.8	(24.8, 28.8)	Oregon	27.3	(25.7, 29.0)
Illinois	28.1	(26.4, 29.9)	Pennsylvania	29.1	(28.1, 30.1)
Indiana	31.4	(30.1, 32.7)	Rhode Island	25.7	(24.1, 27.4)
Iowa	30.4	(29.1, 31.8)	South Carolina	31.6	(30.4, 32.8)
Kansas	29.9	(28.7, 31.0)	South Dakota	28.1	(26.5, 29.8)
Kentucky	31.3	(29.9, 32.6)	Tennessee	31.1	(29.6, 32.7)
Louisiana	34.7	(33.1, 36.4)	Texas	29.2	(27.8, 30.5)
Maine	28.4	(27.2, 29.5)	Utah	24.3	(23.3, 25.3)
Maryland	27.6	(26.3, 28.9)	Vermont	23.7	(22.3, 25.1)
Massachusetts	22.9	(22.0, 23.8)	Virginia	27.4	(26.0, 28.7)
Michigan	31.1	(29.8, 32.3)	Washington	26.8	(25.8, 27.8)
Minnesota	25.7	(24.7, 26.8)	West Virginia	33.8	(32.2, 35.4)
Mississippi	34.6	(33.0, 36.2)	Wisconsin	29.7	(27.8, 31.6)
			Wyoming	24.6	(22.8, 26.4)

plus DC (e.g., by removing Colorado from the data), the medium can be calculated by summing the two middle values (25th Maryland 27.60% and 26th Illinois 28.10%) and dividing them by two. Such logic of taking the proportional value can also be applied to the calculation of the first and third quartile.

Another way to describe continuous data is the *mean* (*sample mean*, M ; *population mean*, μ) which is also called the *arithmetic average*. The mean can be computed by adding all the scores in

a distribution of continuous data and dividing by the total number of scores. In the obesity rate example, in order to calculate the mean, the obesity rate of all 51 locations should be summed up first and the number is 1,425.20. Then, the mean can be obtained by dividing the sum obesity rate by 51, and the value is 27.95. On average, 27.95% of Americans reported themselves as obese. The mean is the most appropriate and effective measure of central tendency of a set of continuous data. However, it is sensitive to every change in

the data and nonresistant to extreme scores. For instance, if six students are supposed to take an exam but one was sick on that day, suppose the grades of those five students who took the exam are 90, 85, 100, 80, and 75, with an average of 86. But if all the six students are taken into account in calculating the mean, it will drop sharply to 71.7, which cannot reflect students' ability in a fair way.

Describing categorical variables is different from describing continuous variables, since rank order has little meaning for categorical variables. One way to describe categorical variables is by *frequency*, the number of occurrences. The score or value in a distribution that occurs most frequently is the *Mode* (*Mo*). The mode is appropriate to describe a set of nominal data, since categorical data are in the form of categories and no mathematical analyses (e.g., computing the mean) can be conducted in a meaningful way. In a nominal data set, the mode points out the category or numerical score that is most "typical" in terms of the number of times it appears. On the contrary, *antimode* is the score or value that appears the least frequently. Take the data set by the CDC as an example. Among the 10 states plus DC with least obesity rate, four are in Northeast, zero are in the Midwest, one are in the South, and five are in the West. So, the mode is the West while the antimode is the Midwest. Therefore, we can know from the data that states in the West are more likely to have a low obesity rate. In the cases that more than one category or score occur the same amount of times, such distribution is called *bimodal* if there are two modes or *multimodal* distribution if three or more modes are present in contrast to a *unimodal* distribution where only one mode exists. For instance, when we look at the 10 least obese locations, the distribution is unimodal, with only one mode, the West. However, the distribution can be bimodal if we only look at the first nine locations with least obesity rate, since both the West and the Midwest occur four times.

Median, mean, and mode are all measure of central tendency but differ in several ways. The mode is considered as the simplest measure of central tendency but is mainly used in describing categorical data. The median is appropriate for describing the center point of a set of quantitative data. One advantage the median has over the mode is that it takes into account the entire set of

scores in a distribution, rather than only one or a few scores. A more important advantage is that median is not swayed by extreme scores, referred to as *outliers*, and considered as a *resistant statistic*. On the flip side, however, the ability of the median to resist the influence exerted by extreme scores is also its primary weakness compared to the mean, as the median is not sensitive to changes in the data set. Therefore, the mean is considered as a *nonresistant statistic* in contrast to the resistant static of the median. There are some types of procedures that can be used to deal with extreme scores. For instance, *trimmed means* are obtained by removing the most extreme scores entirely, which is frequently used in calculating the grades evaluated by a panel of judges to avoid bias, and *Winsorizing* is to change the most extreme scores to the next less extreme scores.

Section 2.02 Measures of Dispersion

Providing central tendency summary statistics is not enough to understand data because measures of central tendency indicate little about how the scores in a distribution vary from this representative score. *Measures of dispersion* or *measures of variability*, usually applied to continuous data, report how much scores vary from each other or how far they are spread around the center point of a data set. Three common measures of dispersion are range, variance, and standard deviation.

Range (or *span*) is calculated by subtracting the minimum from the maximum in a distribution and indicates the distance between the highest and lowest scores in a distribution. Although the range provides a general sense of how much the data spread out across a distribution, a major problem is that it is a *nonresistant measure* and overly sensitive to extreme scores. In other words, one or two outlier can make a distribution look more dispersed than it actually is. To compensate for this problem, we can divide a distribution at the median to calculate the *lowsprad*, which is the range between the median and the minimum, and the *highsprad*, which is the range between the median and the maximum. Another way is to calculate the distance between the first and the third quartile, which yields that *interquartile range* (*IQR*), also referred to as the *quartile deviation* or *mid-spread*. The *IQR* can be further divided to

produce the semi-quartile range if finer gradations are needed.

Variance (sample variance, s^2 ; population variance, σ^2) is an index in squared units of the average distance of scores around the mean in a continuous data distribution. The variance can be calculated by summing the square distance of each value from the mean, and dividing the sum value by the number of values. The formula is $\sigma^2 = (X - \bar{X})^2/N$. The **deviation score** is the amount that one score differs from the mean. To sum up both the positive and negative deviation scores in a meaningful way, *sum of squares* (SS) is always adopted in statistics by adding the squared deviation scores cumulatively to (a) convert negative values into positive ones and (b) to amplify deviations. However, a variance score can be confusing because it is expressed in units of squared deviations about the mean rather than the original trait of measurement. Therefore, *standard deviation* (SD; *sample standard deviation*, s ; *population standard deviation*, σ) is more frequently used, which is a measure of dispersion that explains how much scores in a set of continuous data vary from the mean and is expressed in the original unit of measurement. Standard deviation is found by taking the square root of the variance. The mean and standard deviation are the measures of central tendency and dispersion reported most frequently by researchers when analyzing continuous data.

In statistics, the *standard score* (*standard normal deviates*) is often used to provide a common unit of measurement indicating how far any particular score is away from its mean. Researchers often report how many standard deviations a particular score is above or below the mean of its distribution. The most frequently type of standard score is the z -score, $z = (X - \bar{X})/SD$, which can be calculated by dividing its deviation score by the SD for the distribution. The data kept in their original unit of measurement and not converted to standard scores are *unstandardized* or *raw scores*. To interpret results in a more meaningful way, researchers usually convert raw scores into standard scores.

Article III. Probability and Statistical Inferences

Statistical inference is the process through which people can infer facts about a population using

noisy statistical data where uncertainty must be accounted for. During the inferential process, the sample mean, the standard deviation, and all the data level quantities are estimators. A statistical framework is needed to make those estimations described within the collected data set. In other words, an inference needs to be performed to generalize or use the data set collected to make conclusions beyond. Inferential statistics is mainly used to (a) estimate the characteristics of a population from data collected from a sample and (b) test for significant statistical differences between groups and significant statistical relationships between variables.

Because the size of most populations usually makes it too difficult or impossible to conduct a census, researchers have to study a *sample(s)* that are actually obtained from a targeted population and use the results from that sample to describe the population. Suppose that the Secretary of Education is interested in the GRE scores of all graduate students in the United States and has collected the scores of 100 graduate students. In this case, these 100 scores are the sample, and the scores of all graduate students in the United States is the population to be estimated. The population characteristics are *parameters* and the sample characteristics are *statistics*. The ultimate purpose for researchers is to generalize the results found from a sample back to its parent population. To achieve this, *parametric statistics* are often used as the estimation procedures. The statistics computed are called *estimates* and the formulas used to compute estimates are called *estimators*. Estimates of population parameters from sample statistics are possible when two assumptions are met: (a) the variable(s) of interest is distributed “normally” in the population being studied and (b) a random sample has been selected from that population.

Estimation procedures start with the assumption that scores on the variable(s) of interest in the population being studied are distributed in the shape of a symmetrical bell. This type of distribution is called a *normal distribution* or a *normal/bell-shaped curve*. For a normal distribution, the center point is in the exact middle of the distribution and is the highest point of the curve. The center point is the same as the mean, median, and mode of the distribution. Because of the

symmetrical shape of a normal distribution, researchers know the proportion of scores that falls within any particular area of the curve. In normal distributions, approximately 68% of the data values are within the range from the mean $\pm$ one *SD*. About 95% of data values are within the mean $\pm$ two *SDs*, and 99.7%, almost all of the data values are within three *SDs* of the mean (Figure 1). The curve continues to run on theoretically forever, a curve that runs on, getting ever closer to the line but never touching it is called *asymptotic*. The normal distribution provides information about the relative probability of a score failing in any given area of the curve. Understanding the normal distribution and the percentage of values falling in specific ranges has much implication to the inferential statistics.

Two characteristics are used to describe the distributions that may not be in the shape of a normal curve. First is the *kurtosis*, which refers to how pointed or flat is the shape of a distribution of scores. The shape of the normal curve, with the center point being the highest peak and the two tails leveling off in a similar way, is called *mesokurtic*. However, if a curve is tall and sharply pointed, with scores clustered around the middle, it is called *leptokurtic* or *peaked*. If a curve that is flat and more spread out because scores do not cluster around the middle but are dispersed rather evenly across the distribution, it is referred to as *platykurtic* or *flat* (Figure 2). The other characteristic is *skewness*, indicating that the majority of scores are toward one end of a distribution and the curve tails off across the rest of the distribution, providing an asymmetrical distribution. There are two types of skewness.

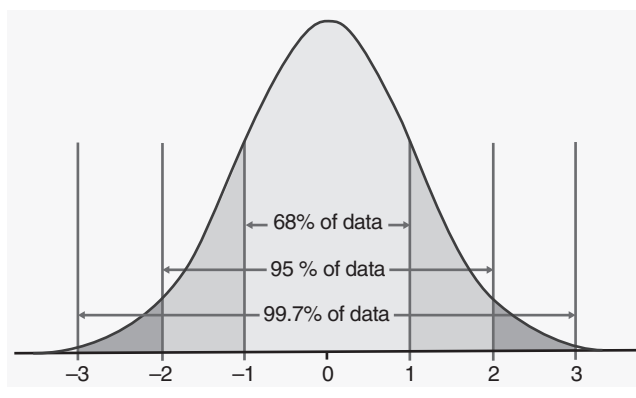


Figure 1 A Normal Distribution

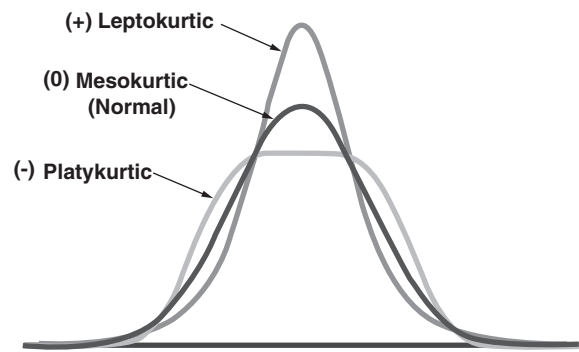


Figure 2 General Terms of Kurtosis

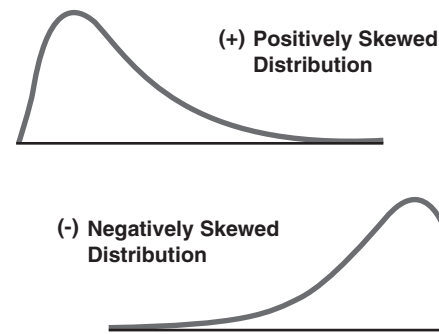


Figure 3 Skewness

The first is *positively skewed* or *right-skewed distribution* which occurs when the tail runs to the right side of the curve. In this case, the mode is less than the median, which is less than the mean. The other type of skewness is the *negatively skewed* or *left-skewed* distribution which occurs when the tail runs to the left side of the curve. For a negatively skewed distribution, the mode is larger than the median which is larger than the mean (Figure 3).

Article IV. Random Sampling

The second assumption that informs estimation procedures is the selection of a random sample from the population being studied. The best assurance of a representative sample is random selection of population members, such that each population member has an equal chance of being selected for the sample. According to the *law of large numbers* (LLN), the average result will converge to the experiment's expected value if

an experiment is repeated for infinite times. A random sample is representative of a population because if a population is normally distributed on a variable, it is probable that a random sample drawn from that population will also be normally distributed on that variable. Another very important principle of inference that is related to sampling is the *central limit theorem* (CLT), which indicates that with any distribution of the original variable, for samples of a sufficiently large size, the real distribution of means is almost approximately normal at most of the times. The CLT addresses two aspects about sampling: (a) the larger samples give more accurate results than do smaller samples and (b) the random samples allow researchers to work under the normal curve. Therefore, random samples are highly concerned in statistical inference. It is always better to go with the random sample if we have a choice. However, if nonrandom samples are used, the normal distribution needs to be checked.

If we draw samples many times from the population, we will get a *sampling distribution*, which occurs when treating the random sample means as a distribution of data, and can also be formed of other estimators, such as percentages or median scores. The mean of sampling distribution means is referred to as the *grand mean*. During sampling, the observed value of an estimator varies from sample to sample. Such variability is called *sampling variability*, which is expected in statistics. Sampling variability occurs because of the randomness of the data. Because of the variability, *standard error of the mean* (SE_m) is applied to tell how much these random sample means are likely to differ from the grand mean of the sampling distribution. Standard error is calculated by dividing the standard deviation by the square root of the sample size, $SE_m = SD/N$.

Since we assume that any characteristic normally distributed in a population will also be normally distributed in a random sample selected from that population, it is possible to make estimation from the sample to the population. However, due to the sampling variability, the estimation is always a range, referred to as the *confidence interval* (CI), which is the range of scores of the random sample means (or other statistics). The

actual values of the upper and lower limits of a confidence interval are referred to as the *confidence limits*.

Confidence interval is highly associated with a *confidence level*, the percentage of random sample means (or other statistic) in a sampling distribution covered by the estimate of a population parameter. In other words, confidence level represents our confidence in estimating that an actual mean from a sample accurately represents the true population mean. Although the confidence level is always arbitrary, a typical cut-off used in social sciences is 95%, with a confidence interval being the mean $\pm 1.96 SE_m$. If we want a more precise range of mean $\pm 1.645 SE_m$, the confidence level drops to 90%, while if we want to be more careful and confident, we can choose 99% confidence level by including the mean $\pm 2.576 SE_m$. On the other hand, $(1 - \text{confidence level})$ is the probability of being wrong, referred to as *alpha* (α) or the *critical value*.

Article V. Statistical Testing

Statistical tests are also referred to as the *hypothesis tests*, or *tests of significance*, which is like a proof by contradiction in mathematics. Starting with an assumption, people will conduct statistical work under that assumption. Statistical tests generally consist of four steps: (a) determine the null hypothesis (H_0) and the alternative hypothesis (H_a); (b) collect the data and calculate a test statistic; (c) determine p value, or how unlikely the test statistics is if the H_0 is true; and (d) make a conclusion based on the p value and the context of the problem.

The first step is to formulate a hypothesis to test and specify an alternative to that hypothesis. Statistical testing is analogous to a court case, that a defendant is regarded as innocent until proven guilty. In statistics, the presumption of innocence corresponds to the *null hypothesis* (H_0). The H_0 often argues that nothing is happening, such as there is no relationship between the variables or there is no difference between the two groups. The H_0 is taken as a working hypothesis and is assumed to be true until it is disproved by the data. However, what the researchers really want to know is whether the opposite is true, by bringing up an *alternative hypothesis* (H_a or H_1), also called as the

research hypothesis. Just like the prosecutor who collects evidence to prove that the defendant is guilty, we conduct hypothesis testing to prove that H_0 is wrong and H_a is true, by seeing how likely it is that the H_0 happens by chance. If a direction of an outcome can be predicted, a *one-sided* H_a can be used, by stating that the outcome is either greater or less than 0. Otherwise, a *two-sided* H_a should be applied, by arguing that the outcome is not equal to 0. Suppose a public health researcher wants to evaluate the effect of a smoking intervention by conducting statistical testing. If the researcher knows that the intervention cannot make the tobacco use worse from theory or former research, they can use one-sided H_a . Otherwise, the researcher has to stick to the two sided H_a , if no clue of the intervention effect is available. In most situations, two-sided tests are appropriate.

So, the second step is to accumulate some evidence. Similar to the evidence collected by the police and prosecution in a court case, in statistics, we need to put together the evidence by analyzing our data. In order to make the data useful, we have to summarize the data into the *test statistic*, which is a numerical summary of the data. A test statistic is constructed assuming that H_0 is true. Then, we need to collect some data and summarize them into a test statistic.

Once the evidence has been collected and summarized for the judge and jury, the third step is to determine how unlikely the test statistic is if the H_0 is true with p value. p Value, a number between 0 and 1, transforms a test statistic into a probabilistic scale and quantifies how strong the evidence is against the null hypothesis. To decide whether the results could just happen due to chance, we wonder if H_0 is true, how likely it would be to observe a test statistic, or summary of our data of this magnitude just by chance. The numerical answer to this question is the p value. The smaller the p value, the more unlikely it is that we assume the H_0 is true, and the stronger the evidence we have against the H_0 . It is noteworthy that the p value is *not* a measurement of how likely it is that the H_0 is true. We can't put a probability value on H_0 , because it's not random. What a p value can tell is how likely the observed data would be, if an H_0 is true. It's a measure of the strength of the evidence against the H_0 .

So finally, we need to make a conclusion of whether the summary of our data of this magnitude is just by chance. In statistics, if the p value isn't small, we do not have strong enough evidence to support the H_a and have to conclude that our data are consistent with the H_0 . We can't conclude H_0 is exactly true, but we can't reject it either. Alternatively, if the p value is small, which means that we have data that are unusual enough, we have evidence to rule out the possibility that the H_0 is true, and we got the observed data by chance. In such case with a small p value, we conclude that we have sufficient evidence to reject the H_0 in favor of H_a .

Qinghua Yang

See also Modeling; Statistical Inference, Bayesian; Statistical Inference, Frequentist

Further Readings

- Cowles, M. (1989). *Statistics in psychology: An historical perspective*. Lawrence Erlbaum.
- Field, A. (2009). *Discovering statistics using SPSS* (3rd ed.). Sage.
- Kolenikov, S., Steinley, D., & Thombs, L. (Eds.). (2010). *Statistics in the social sciences: Current methodological developments*. Wiley.
- Mendenhall, W., Beaver, R., & Beaver, B. M. (2020). *Introduction to probability and statistics*. (15th ed.). Cengage Learning.
- Stevens, J. P. (2012). *Applied multivariate statistics for the social sciences*. Routledge.
- Vogt, W. (1999). *Dictionary of statistics and methodology: A nontechnical guide for the social sciences* (2nd ed.). Sage.
- Weisberg, H. (1992). *Central tendency and variability*. Sage. <https://dx.doi.org/10.4135/9781412983617>

STATISTICS, COMPLETENESS IN

Completeness is a probabilistic property that bestows a form of uniqueness on a statistic $h(X)$ in the context of a statistical model $M\theta(x)$ and crystallizes the link between sufficient and ancillary statistics.

Ronald A. Fisher's (1922) path-breaking paper recast Karl Pearson's descriptive statistics (Yule, 1916) into a model-based statistical induction grounded on the concept of a parametric statistical model whose generic form is:

$$M\theta(x) = \{f(x; \theta), \theta \in \Theta\}, x \in RX_n, \text{ for } \Theta \subset R_m, m < n, \quad (1)$$

where $R := (-\infty, \infty)$, $f(x; \theta)$ denotes the (joint) *distribution of the sample* $X := (X_1, \dots, X_n)$, RX_n the sample space, and Θ the parameter space. The likelihood function is defined by: $L(\theta; x_0) \propto f(x_0; \theta)$, $\theta \in \Theta$.

Example 1. The *simple Normal (N) model* is specified by:

$$\begin{aligned} X_t &\sim \text{NIID}(\mu, \sigma^2), \theta := (\mu, \sigma^2) \in \Theta := R \times R_+, \\ x_t &\in R, t \in N := (1, 2, \dots, n, \dots), \mu = E(X_t), \\ \sigma^2 &= \text{Var}(X_t), f_N(x; \theta) = (2\pi\sigma^2)^{-n} \exp - 1/2\sigma^2 \sum_{t=1}^n (x_t - \mu)^2. \end{aligned} \quad (2)$$

Example 2. The *simple Bernoulli (Ber) model* is specified by:

$$\begin{aligned} X_k &\sim \text{BerIID}(\theta, \theta(1 - \theta)), \theta \in \Theta := (0, 1), \\ x_k &= 0, 1, k \in N, \theta = E(X_t), \text{Var}(X_t) = \theta(1 - \theta), \\ f_B(x; \theta) &= \theta^n x^n (1 - \theta)^{n(1 - x)}. \end{aligned} \quad (3)$$

where $X = (1/n)k = 1/n X_k$.

Frequentist inference revolves around Borel measurable (well-behaved) mappings of the form:

$$h(\cdot) : RX_n \rightarrow R_k, n > k \geq 1,$$

Fisher (1922) named *statistics* $Y = h(X)$, such as estimators, confidence interval (CI) bounds, test statistics, and predictors, whose sampling distributions are derived via:

$$F(y) = P(Y \leq y) = \int \{x: h(x) \leq y\} f(x; \theta) dx, \quad \forall y \in R_k, \quad (4)$$

$f(y; \theta) = \partial F_n(y)/\partial y$, $\forall y \in R_k$, and provide the foundation of model-based frequentist inference. Of particular interest in this context are the sufficient and the ancillary statistics, introduced by Fisher in the early 1920s (see Lehmann & Scholz, 1992).

Sufficient. In the context of a statistical model $M\theta(x)$, a statistic $S := S(X)$ is said to be *sufficient* for θ if:

$$f(x | s; \theta) = f(x | s), \forall x \in RX_n, \quad (5)$$

or equivalently (see Cox & Hinkley, 1974, p. 22):

$$f(x; \theta) = f(x | s) \cdot f(s; \theta), \forall x \in RX_n, \forall s \in R_k. \quad (6)$$

That is, the distribution of the sample $f(x; \theta)$ can be separated into a product of:

- (i) a conditional distribution of X given $S = s$ that is free of θ and
- (ii) the marginal distribution of S which depends on θ .

The sufficient statistic S is said to induce the family of sampling distributions $\{f(s; \theta), \theta \in \Theta, s \in R_k\}$.

The statistic S is said to be *minimal sufficient* for θ , if for any other sufficient statistic $T := T(X)$ there exists a function $g(\cdot)$ such that $S = g(T(X))$; in practice S has the lowest dimension possible.

Ancillary. In the context of a statistical model $M\theta(x)$, a statistic $V := V(X)$ is said to be *ancillary* for θ , if its sampling distribution:

$$f(v), v \in R_{v,k}, n > k \geq 1,$$

is free of (does not depend on) θ (see Lehmann and Romano, 2005).

Intuitively, an ancillary statistic V contains *no information* about θ . The statistic V is said to be *maximal ancillary* for θ , if for any other ancillary statistic $U(X)$ there exists a function $g(\cdot)$ such that $U(X) = g(V(X))$; a maximal ancillary statistic V has the highest possible dimension k .

As argued by Erich Leo Lehmann and Fritz Scholz,

Ancillarity is in a certain sense the dual of sufficiency. If S is a sufficient statistic, then any inference can be based solely on S , and the conditional distribution of the full data set X given S is independent of its parameters. Conversely, if V is ancillary, inference may be based entirely on the conditional distribution of X given V , while the distribution of V is independent of the parameters. In this duality, a maximal ancillary corresponds to a minimal sufficient statistic. They differ however in that a minimal sufficient statistic is essentially unique and that explicit methods for its construction are available, neither of which is the case for maximal ancillaries. (1992, p. 33)

In an attempt to narrow down the class of possible maximal ancillary statistics, Lehmann and Scholz (1992) proposed restricting such statistics to those that are functions of the minimal sufficient statistics, that is, $V(X) = V(h(S))$.

Completeness, introduced by Lehmann and Henry Scheffé (1950), provides the key to elucidating the connection between sufficiency and ancillarity.

Completeness. The family of distributions defining a statistical model $M_\theta(x)$ is said to be *complete* if for any statistic $h(X)$ such that $E(h(X)) = 0$ for all $\theta \in \Theta$, that is, its mean is free of θ , implies that:

$$P_{h(X)=0} = 1, \text{ for all } \theta \in \Theta, \quad (7)$$

that is, $h(x) = 0$ with probability 1. If the Property (7) holds only for all bounded functions $h(\cdot)$, then the family is said to be *boundedly complete*. The intuition underlying (7) is that the only unbiased estimator of $0 \in \Theta$ is the zero function $h(X) = 0$. The same definition also applies to the sampling distribution $\{f(s; \theta), \theta \in \Theta, s \in R_k\}$ induced by a statistic $S := S(X)$.

For a sufficient statistics $S := S(X)$, completeness secures the uniqueness of certain properties based on a sufficient statistic S , such as unbiasedness and minimum variance.

Lehmann–Scheffé theorem. If there is a complete sufficient statistic $S := S(X)$ for θ in the context of a statistical model $M_\theta(x)$, then there is a unique unbiased estimator $\theta = h(S)$ of θ .

Intuitively, a complete sufficient statistic S contains no *ancillary* information in the sense that $g(S)$ is an ancillary statistic if and only if for some constant c , $P[g(S) = c] = 1$ for all $\theta \in \Theta$ (see Knight, 2000). This suggests that a way to construct ancillary statistics is to seek functions of the form $g(S)$ whose distribution is constant for all values of θ (free of θ), such as the residuals from the estimated model as argued next.

Sufficiency, Ancillarity, and Completeness

The amalgam of completeness, sufficiency, and ancillarity has given rise to several important results in statistical modeling and inference, including the following.

Lemma 1. If $S(X)$ is a *complete* sufficient statistics for θ , in the context of a statistical model $M_\theta(x)$, then it is a *minimal* sufficient statistic (see Lehmann & Scheffé, 1950).

Example 1 (continued). The simple Normal model in (2), in addition to having $S(X) = X_n = 1/n \sum_{i=1}^n X_i$, $s^2 = 1/n \sum_{i=1}^n (X_i - X_n)^2$ being a minimal sufficient statistic, $f(x; \theta)$ and $f(s; \theta)$ define families of distributions which are complete (see Lehmann & Romano, 2005).

Example 2 (continued). The simple Bernoulli model in (3), in addition to having $S(X) = X$ as a minimal sufficient statistic, $f(x; \theta)$ and $f(s; \theta)$ define families of distributions which are complete (see Lehmann & Romano, 2005).

Sufficiency principle (SP). If $S(X)$ is a sufficient statistic for θ , then any inference about θ should depend on the sample X only through $S(X)$. That is, exclusive reliance on $f(s; \theta)$ for inference purposes forfeits no relevant sample information since (6) implies that $f(x; \theta) \propto f(s; \theta)$, $\forall s \in R_k$.

A very interesting special case of the reduction in (12) arises when sufficiency and ancillarity arise together in the context of a complete family of distributions defining an ancillary statistic $V(X)$ and a complete minimal sufficient statistic $S(X)$ for θ .

Basu's theorem. For a statistical model $M_\theta(x)$, when $S(X)$ is a *complete sufficient* statistic, then it is *independent* of every *ancillary statistic* $V(X)$. Equivalently, $S(X)$ is independent of any statistic $V(X)$ whose distribution is free of (does not depend on) θ (see Basu, 1955).

This is a remarkable result which crystallizes the connection between sufficiency, ancillarity, and independence and has numerous applications in statistic (see Lehmann and Romano, 2005).

Basu's theorem has its most wide applicability in the context of the *exponential family* of distributions, introduced by Fisher (1934), which is complete and admits minimal sufficient statistics.

Exponential family. The *regular canonical form* of a p -parameter exponential family density is:

$$f(x; \theta) = a(\theta)h(x)\exp\theta^T s(x), \quad x \in R_X^m, \quad \theta \in \Theta \subset R_p, \quad (8)$$

$$a(\theta) = [1/c(\theta)], \quad c(\theta) = \int_{R_X^m} h(x)\exp\theta^T s(x)dx > 0, \quad \theta^T s(x) = \sum_{i=1}^p \theta_i s_i(x),$$

Θ is an open, convex subset of R_p , $p = \dim(S) = \dim(\theta)$, where “dim” denotes “dimension.”

The most important property of the family in (8) is that $S := S(X)$ is a minimal sufficient statistic whose sampling distribution also belongs to the same family:

$$f(s; \theta) = a(\theta)g(s)\exp\theta^T s, s \in R_p, \theta \in \Theta \subset R_{xp},$$

where $g(s) = \{s(x) = s\}h(x)$ when x is discrete and $g(s) = \{x: s(x) = s\}h(x)dx$ when x is continuous (see Brown, 1986).

An important property of the exponential family is that the joint distribution of $x_1, x_2, \dots, x_n$ also belongs to the exponential family:

$$f(x_1, x_2, \dots, x_n; \theta) = a(\theta)n\exp\theta^T i = 1n s(x_i)_i = 1n h(x_i), \forall x_i \in R_m, \theta \in \Theta \subset R_p.$$

The exponential family includes distributions such as Normal, inverse-Normal, Log-Normal, exponential, gamma, chi-square, beta, Dirichlet, Wishart, inverse Wishart, Rayleigh, Bernoulli (univariate, multivariate), Poisson, geometric, binomial (univariate and multivariate, with fixed number of trials), negative binomial (with fixed number of failures), and logarithmic. It excludes other known distributions, such as the Cauchy and the Student's t .

Complete Sufficiency and Similar Tests

Let the parameters of a statistical model $M\theta(x)$ be $\theta := (\phi, \psi) \in \Phi \times \Psi$, and $S := (S_1, S_2)$ is a sufficient statistic for $\theta := (\phi, \psi)$, where S_1 is sufficient for ϕ and S_2 is sufficient for ψ , that is,

$$f(s; \phi, \psi) = f(s_1|s_2; \phi) \cdot f(s_2; \psi), \quad \forall (\phi, \psi) \in \Phi \times \Psi. \quad (9)$$

For the hypotheses $H_0: \phi \in \Phi_0$ versus $H_1: \phi \in \Phi_1$, let a Neyman–Pearson (N-P) test be $T\alpha := \{d(X), C_1(\alpha)\}$, where $d(X)$ is a test statistic and $C_1(\alpha) = \{x: d(x) > c\alpha\}$ a rejection region. In general, the Type I error probability and the power of $T\alpha$ are functions of $\psi \in \Psi$ defined by:

$$\max_{\phi \in \Phi_0} P(d(X) > c\alpha; H_0) = \alpha(\psi), P(d(X) > c\alpha; H_1) = P(\phi, \psi), \forall \phi \in \Phi_1, \forall \psi \in \Psi,$$

specifying a *non-similar test* whose error probabilities are functions of ψ . However, when S is a *complete* sufficient statistic for $\theta := (\phi, \psi)$, one

can condition on $S_2 = s_2$ to render the Type I error probability and the power of $T\alpha$ free of ψ :

$$\max_{\phi \in \Phi_0} P(d(X) > c\alpha | S_2 = s_2; H_0) = \alpha, \\ P(d(X) > c\alpha | S_2 = s_2; H_1) = P(\phi), \forall \phi \in \Phi_1,$$

specifying a *similar test*. Statistical models with such complete sufficient statistics satisfying (9) are said to have a *Neyman structure* (see Lehmann & Romano, 2005).

Modeling Versus Inference: A Case for Ancillarity

Another important result that follows from a combination of a complete sufficient statistic $S(X)$ and a maximal ancillary statistic $V(X)$ for $\theta \in \Theta \subset R_p$ in the context of a statistical model $M\theta(x)$.

Consider a one-to-one mapping from the sample X into two vector statistics $S(X)$ and $V(X)$:

$$X \rightarrow (S(X), V(X)), X (n \times 1), S(X) (p \times 1), \\ V(X) ((n - p) \times 1), \quad (10)$$

$$\dim X = \dim V + \dim S, \\ J = \det \partial s \partial x : \partial v \partial x > 0 \quad (11)$$

In the general case where *no restrictions* are imposed on (S, V) , $f(x; \theta)$ can be separated into:

$$f(x; \theta) = J \cdot f(s, v; \theta) = J \cdot f(s|v; \theta) \cdot f(v; \theta) = J \cdot f(v|s; \theta) \cdot f(s; \theta), \forall s, v \in R_{sp} \times R_{vn-p}. \quad (12)$$

What is intriguing about (12) is how it simplifies when one assumes that:

(a) $S := S(X) = (S_1, \dots, S_p)$ is a *complete (minimal) sufficient* statistic,

(b) $V := V(X) = (V_{p+1}, \dots, V_n)$ a *maximal ancillary* statistic. Basu's theorem implies that:

$$f(x; \theta) = J \cdot f(s; \theta) \cdot f(v), \forall s, v \in R_{sm} \times R_{vn-m}, \quad (13)$$

and (c) $S(X)$ and $V(X)$ are *independent*.

The clear separation of $f(s; \theta)$ and $f(v)$ in (13) stemming from (c) implies that inference can be based exclusively on $f(s; \theta)$ (SP) since the likelihood function reduces to $L(\theta; x_0) = c(x_0) \cdot f(s; \theta)$, $\forall \theta \in \Theta$, where $f(v)$ becomes a component of the proportionality constant $c(x_0)$. On the other hand, since $f(v)$ is free of θ , it can be used to validate $M\theta(x)$ using misspecification testing since the error probabilities in testing its validity do not depend on θ (see Spanos, 2019).

The p-parameter exponential family of distributions is tailor-made for this since it provides a general framework where one can separate the inference from the modeling facets.

Example 1 (continued). For model evaluation purposes the simple Normal model in (2), we know that $S = (X_n, s_2)$ is a minimal sufficient statistic for $\theta = (\mu, \sigma^2)$. Not surprisingly, the maximal ancillary statistics relates to the *residuals*:

$$u_t = (X_t - X_n) \sim N(0, (n-1)\sigma^2), t \in N,$$

whose distribution depends on σ^2 . However, we know that:

(i) the residuals $u_1, u_2, \dots, u_n$ are independent of X_n , (ii) X_n and s_2 are independent, and thus Basu's theorem implies that (iii) the *studentized* residuals $v := v_1, v_2, \dots, v_n$:

$$v_t = (X_t - X_n) / s_2 \sim St(n-1), t = 1, 2, \dots, n,$$

are independent of X and s_2 . Hence, v constitutes an ancillary statistic since its distribution is free of (μ, σ^2) , which is also a function of the minimal sufficient statistic X, s_2 (see Lehmann and Romano, 2005). This renders the studentized residual statistic v ideal for model validation purposes. Given that the distribution of $v_1, v_2, \dots, v_n$ are singular, $S := (X, s_n)$ and $V := (v_3, \dots, v_n)$ constitute the relevant minimal sufficient and maximal ancillary statistics, respectively; $\text{rank}(X) = n$, $\text{rank}(S) = 2$, and $\text{rank}(V) = n - 2$. As argued by Spanos (2019, chap. 15), the residuals $\{(v_t, v_{t+2}), t = 3, 4, \dots, n\}$ can be used to construct joint misspecifications tests based on auxiliary regressions to ascertain the validity of the model $[M\theta(x)]$ assumptions, for many statistical models.

Aris Spanos

See also Statistics; Statistical Inference, Bayesian; Statistical Inference, Frequentist

Further Readings

- Basu, D. (1955). On statistics independent of a complete sufficient statistic. *Sankhya*, 15, 377–380.
Brown, L. D. (1986). *Fundamentals of statistical exponential families: with applications in statistical decision theory* (IMS Lecture notes and

monograph series 9). Institute of Mathematical Statistics.

- Cox, D. R. & Hinkley, D. V. (1974). *Theoretical statistics*. Chapman & Hall.
Fisher, R. A. (1922). On the mathematical foundations of theoretical statistics. *Philosophical Transactions of the Royal Society A*, 222, 309–368.
Fisher, R. A. (1934). Two new properties of mathematical likelihood. *Proceedings of the Royal Society, Series A*, 144, 285–307.
Knight, K. (2000). *Mathematical statistics*. Chapman & Hall/CRC.
Lehmann, E. L., & J.P. Romano, J. P. (2005). *Testing statistical hypotheses* (3rd ed.). Wiley.
Lehmann, E. L., & Scheffé, H. (1950). Completeness, similar regions and unbiased estimation I. *Sankhya*, 10, 305–340.
Lehmann, E. L., & Scholz, F. W. (1992). Ancillarity. In M. Ghosh & P. K. Pathak (Eds.), *Current issues in statistical inference: Essays in honor of D. Basu* (pp. 32–51) (IMS Lecture Notes and Monograph Series 17). Institute of Mathematical Statistics.
Spanos, A. (2019). *Probability theory and statistical inference: Empirical modeling with observational data* (2nd ed.). Cambridge University Press.
Yule, G. U. (1916). *An introduction to the theory of statistics* (3rd ed.). Griffin.

STRUCTURALISM (INTRODUCTION TO THEORY)

History is replete with views and movements sharing the same name, though exhibiting considerable divergence in detail. The term *structuralism* is no exception. Structuralisms have emerged in fields as diverse as anthropology, architecture, economics, linguistics, literary criticism, mathematics, philosophy, psychology, and sociology. What they all have in common is an emphasis on structural analysis—that is, on understanding a given system by paying close (and sometimes exclusive) attention to the relations that organize its elements. Even within philosophy, structuralist views have popped up in a number of subfields including aesthetics, the philosophy of language, the philosophy of mathematics, and the philosophy of science. This entry focuses on two prominent and interrelated structuralisms in the

philosophy of science, namely structuralism in the scientific realism debate and structuralism in the scientific theory and representation debates.

Structuralism in the Scientific Realism Debate

The scientific realism debate is primarily a discussion over what epistemological (sometimes also over what metaphysical and axiological) commitments one should endorse vis-à-vis scientific theories and models. Roughly speaking, scientific realists—henceforth just *realists*—hold a generally positive epistemic attitude toward theories and models. In more detail, they claim that our best—our most explanatorily and empirically successful—theories and models are likely to be approximately true or to at least contain some nonnegligible truth content in relation to both observable and unobservable features of the world. By contrast (and equally roughly), scientific anti-realists—henceforth just *anti-realists*—hold a generally negative epistemic attitude toward theories and models. That is, they claim that theories and models, even the best ones, are unlikely to contain nonnegligible truth content (and hence are unlikely to be approximately true) in relation to both observable and unobservable features of the world or that we are never in an epistemically secure enough position to assert that they do.

It is worth dwelling a bit on the distinction between observable and unobservable features of the world. Crudely put, observable features are those that are directly detectable through our unaided sensory organs. Unobservable ones are those that resist such detection. Thus, animals, volcanoes, and dwarf planets are observable but viruses, atoms, and subatomic particles are not. In effect, anti-realists are skeptics about many of the posits of modern science because drawing ontological conclusions from instruments typically involves long chains of inferences, each of whose links can be questioned. For example, the detection of subatomic particles relies on circuitous inferences that begin with the observable tracks they presumably leave behind in particle accelerator chambers. The significance of the observable–unobservable distinction has been questioned by some realists. Generally speaking, the realists are

happy to concede the distinction but argue that instruments nonetheless provide some epistemic access to some posits.

Structuralism in the scientific realism debate is a view that has developed against the backdrop of several renditions of the debate. It is more commonly known as *structural realism*, although, as we shall soon see, not all views under its banner are realist. Structural realists argue that our best scientific theories and models have no trouble providing accurate descriptions of various features of the observable world but are limited in their ability to accurately describe the unobservable world. More formally, such theories and models describe the unobservable world up to isomorphism (a technical notion which means, roughly, having the same form or structure). On this view, we can only hope to describe the structure of the unobservable world. One of the earliest motivations for the view appears in the work of Henri Poincaré (1854–1912) and is historical in nature. Crudely put, we can only know the structure of the unobservable world because when we examine the history of science, we discover that any description seeking to go beyond structure comes undone at the next scientific revolution. Structure, contrariwise, seems to survive in at least some approximate form. To give some examples, although the ontological posits in classical descriptions of thermal, optical, electrical, and magnetic phenomena have been supplanted by distinct ontological posits in nonclassical ones, parts of the structure of those descriptions have remained invariant. Let's take the case of optical phenomena. We no longer believe in light as a vibration in an all-pervading elastic medium—what was once known as the *luminiferous ether*. Even so, we use some of the same mathematical structures, such as Henri Fresnel's equations, to predict and explain the corresponding phenomena. To sum up, despite having to discard ontological posits such as the ether, some of the same mathematical structures remain with us. This gives some *prima facie* reason to believe in structural realism. More broadly, if the history of science does indeed follow this pattern, we have good grounds for endorsing structural realism over any other form of realism or anti-realism. (Other cases of structural

continuity have also been proposed—see, e.g., Votsis & Schurz, 2012.)

There are more than a few structural realist views out there. As we cannot explore all of them here, we will restrict our attention to the general divide between the epistemic and ontic families. What we have said so far chimes well with both families. The main difference between them is that the ontic structural realists employ revisionary ontological considerations to reach the same epistemic conclusions (see French, 2014, for details). To be more precise, they argue that the reason why we cannot know anything beyond the structure of the unobservable world is effectively because only structure matters. This claim has been interpreted in distinct ways. In the extreme, it is taken to mean that only structure exists. Objects are just placeholders for the intersections of relations and thus do not strictly speaking exist but are at best useful fictions. A more moderate reading allows for the existence of objects but denies that such entities are individuals and hence demotes their overall role in understanding the unobservable world. Motivations for ontic structural realist views vary, but they are typically drawn from quantum physical considerations. Those who advocate the moderate reading, for example, point out that fundamental objects in quantum physics lack individuality because they fail Leibniz's principle of the identity of indiscernibles, which holds that two objects sharing all their properties are in fact identical.

As alluded to earlier, not all structuralists in the scientific realism debate are realists. There are also structural anti-realists, though they often go by more affirming names like *structural empiricists* or *empiricist structuralists*. The essence of such views is that even though we must remain agnostic with respect to the exact nature or denizens of the unobservable world, our best theories still give us a pretty good description of the observable world. In particular, the claim is made that, if we understand theories as collections of structures, then theories can be used to adequately represent only structures in the observable world. Of note here is that this conception of theories as collections of structures is shared by some of the structural realists, such as Steven French. It is also a

good segue into other types of structuralism, which we will explore in the ensuing section.

Objections to structuralist views abound (for an in-depth discussion of arguments for and against structural realism, see Frigg & Votsis, 2011). Some are more serious than others. Structural realism has been accused of being either too weak to be realist (and hence collapsing into empiricism) or too undistinctive (and hence collapsing into traditional versions of realism). The latter accusation affects also structural anti-realists, their view presumably collapsing into traditional versions of anti-realism. Ontic structural realists have been accused of being incoherent—this is predominantly an objection to the extreme version of the view outlined earlier, since it seems to assert that there can be relations without relata. More generally, all structuralist views in the scientific realism debate seem to be affected by the following questions: (1) What formal frameworks need to be in place to do justice to the requisite notion of structure? and (2) What ontological category plays the role of the non-structural features of the world? As to the first question, the proposals range from second-order logic (and particularly the device known as the *Ramsey sentence*) to set theory, quasi-set theory, and category theory. As to the second question, the proposals range from qualities to intrinsic properties, monadic properties, and haecceities. It is worth noting that this objection runs deeper than the others and has been justifiably underscored by critics.

Structuralism in the Scientific Theory and Representation Debates

The scientific representation debate is primarily concerned with the question of how scientific theories represent their target systems. Taking a stance in this debate often means taking a stance in the debate over the nature of scientific theories. The two main stances in the latter debate are the syntactic view and the semantic view. (A third view, known as the *pragmatic view*, seems to be emerging in popularity, though it is not currently systematic enough or sufficiently distinct from pragmatic-oriented versions of the other two views to deserve attention here.) The *syntactic view* is taken to construe theories as sets of

sentences axiomatized in some meta-mathematical framework, for instance, first-order predicate logic with set theory. By contrast, the *semantic view* is taken to construe theories as sets of models, where a model is a set-theoretic structure. Thus, depending on which view of theories one advocates, the representation relation seems to manifest itself in distinct ways: Target systems are represented by linguistic entities—that is, by sentences—if one follows the syntactic view, or by extralinguistic entities (in other words, abstract set-theoretical structures) if one follows the semantic view.

It is tempting to assume that the semantic view alone deserves the appellation *structuralism* in the scientific theory and representation debates and, additionally, that it alone aligns with structuralism in the scientific realism debate. After all, the semantic view typically construes the representation relation as a mathematical morphism, for example an isomorphism, between theories and the world. Moreover, it explicitly conceives of theories in terms of set-theoretic structures, and such structures, as we have seen, offer one of the ways through which structuralists in the scientific realism debate express the commitments of their views. Such an assumption seems ill-advised. One reason to resist it stems from the fact that some versions of structural realism, namely Ramsey-style structural realism (see Worrall, 1989), are married to the syntactic view of theories. Named after its inventor, the philosopher-mathematician Frank Ramsey, Ramseyfication is achieved by first expressing theories in predicate logic and then logically weakening them by stripping away the interpretations of all non-logical terms meant to denote unobservables. (It is worth noting that the process leaves intact the interpretations of non-logical terms intended to denote the observables. Ramseyfied theories thus restrict what features of the unobservable world can be adequately represented to their logical structure.

Another reason why we should not be tempted by the foregoing assumption stems from the fact that the two views have been variably interpreted. Once we put aside naive readings of each view, the differences between them appear to be few and superficial, if not altogether absent. For example, the syntactic view need not take the

sentences which make up theories to be uninterpreted. Indeed, for each such set of sentences, there is in fact a corresponding set of models which makes those sentences true. Conversely, the semantic view need not take the models which make up theories to be detached from sentences. After all, the function of models, construed as set-theoretical structures, is to make sentences true. (In fact, that's what we mean when we say that models satisfy a theory. The notion of *satisfaction* is central to any discussion of models in the model-theoretic sense.) Other objections that seek to drive a wedge between the two views can be, and have been, questioned on the basis of different interpretations. These include the reputed requirements (attributed to the syntactic view) that theories should be axiomatized and the metamathematical framework should be first-order predicate logic—see Lutz (2012) for a counter—and the reputed requirement (attributed to the syntactic view) that only mathematical (vs. metamathematical) frameworks must be used in analyzing theories and in representing target systems. To summarize, when offered more sophisticated construals, the semantic and the syntactic views seem equally qualified to earn the appellation *structuralist* in their conception of theories and of the representation relation. Both also appear to be a good fit with structuralism in the scientific realism debate.

Thus far, we have presupposed that the representation relation is a kind of mapping between the structures of theories and the structures of the world. There are, however, other accounts of representation which may not be as amiable to the structuralist ideology. Indeed, such accounts purposely try to dissociate themselves from that ideology. On the similarity account, a theory or model represents a target system just in case it is similar to that target system in some relevant respects. Which respects are relevant is sometimes deliberately left unspecified to deal with the kinds of cases that the structuralists are accused of being incapable of handling, namely cases where no direct mapping can be established between theories and the world. The trouble with this approach is that, unless its advocates provide some sort of principled way to circumscribe the similarity

relation, they risk trivializing the view. After all, and as countless philosophers have pointed out, everything is similar to everything in some respect or other. A more promising approach to the similarity account is to insist that the relevant respects are context dependent. Another notable account of the representation relation is *inferentialism*. On this account, a necessary condition for a *surrogate* (the entity that does the representing) to represent a target system is that we can draw inferences from the former to the latter. Inferentialism prides itself in being neutral with respect to the kinds of entities that can be used as surrogates. They can be theories; models; physical, abstract, or fictional objects; and so on. Just like in the case of the similarity account, this strength of the inferentialist account also proves to be a potential weakness. It is always possible to draw some inferences from one object or system to another. Unless the permissible inferences are restricted, the inferentialists risk trivializing their view.

The success or failure of the similarity and inferentialist accounts depends on what kind of information is vital for a representation to genuinely tell us something about target systems. If the requisite kind of information must always be structural in nature, then the two views are faced with an unpalatable dilemma. Either construe the representation relation in a nonstructural way and therefore end up missing the mark or construe it in a structural way and end up collapsing into structuralism.

Conclusion

Whether it is a view within the scientific realism, the scientific theory or the scientific representation debates, structuralism has proved to be highly influential and resilient. Those who wish to study the epistemological, metaphysical, and methodological aspects of these debates are unlikely to make any headway unless they pay such views the attention they deserve.

Ioannis Votsis

See also Metaphysics; Modeling; Scientific Realism; Philosophy of Science

Further Readings

- Cartwright, N. (2019). *Nature, the artful modeler: Lectures on laws, science, how nature arranges the world and how we can arrange it better*. Open Court.
- French, S. (2014). *The structure of the world: Metaphysics and representation*. Oxford University Press.
- Frigg, R., & Votsis, I. (2011). Everything you always wanted to know about structural realism but were afraid to ask. *European Journal for Philosophy of Science*, 1(2), 227–276.
- Lutz, S. (2012). On a straw man in the philosophy of science: A defense of the received view. *HOPUS: The Journal of the International Society for the History of Philosophy of Science*, 2(1), 77–120.
- Poincaré, H. (1952). *Science and hypothesis*. Dover (Original work published 1905).
- Suppes, P. (2002). *Representation and invariance of scientific structures*. CSLI Publications.
- Votsis, I., & Schurz, G. (2012). A frame-theoretic analysis of two rival conceptions of heat. *Studies in History and Philosophy of Science*, 43(1), 105–114.
- Worrall, J. (1989). Structural realism: The best of both worlds? *Dialectica*, (1–2), 99–124.

SYNTAX (INTRODUCTION TO THEORY)

Syntax is the set of rules and principles that govern the structures of sentences; the word also refers to the study of these rules and principles and is thus one of the branches of linguistic knowledge, alongside phonetics, phonology, morphology, semantics, and discourse. Sentences are the basic unit of syntactic analysis; they can be simple, compound, or complex according to their structural composition. The complexity of predication and coherence between the constituents defines the complexity of a sentence. Accordingly, sentences may range from one-word syntactic units to paragraph-long complex structures. Traditionally, the main feature of a sentence has been the expression of one complete thought through logically organized and coherently connected words. As a discipline, however, modern syntax conceives the sentence as a unit consisting of subject and predicate, since a thought can be expressed by a single word, which may not be a sentence.

The term *syntax* derives from the Greek word *syntaxis* meaning *arrangement*. The words are arranged in an organized and structured manner to produce a single logical statement, thought, exclamation, or question. The arrangement of words reveals the relationships of meanings within one syntactic structure. The relationship is coherent, with cohesively bound concepts, meanings or lexico-phraseological units. Syntactic studies focus on sentence structure and connections within and between sentences. Building a syntactic tree is the main basis of syntactic analysis. It displays relations and hierarchies between sentence constituents most vividly.

Therefore, the sentence is the largest unit to which the syntactic rules apply. It is an independent linguistic unit. However, sometimes a sentence omits a part of its structure, for instance, when producing elliptical sentences. In this case, the sentence becomes dependent on the previous or following sentences. Moreover, the production of a complete paragraph requires the logical, cohesive, and coherent connection between sentences that form a completed sequence of thoughts, ideas, and statements expressed in one paragraph.

Approaches to Syntactic Analysis

The fundamental feature of any given natural language is to distinguish logically constructed sentences from illogical sequences of words that are not sentences. Thus, linguistic analysis is concerned with identifying *grammatical* and *ungrammatical* sequences of phonemes, lexical, or grammatical units. By grammatical sequences, it is implied that phonemic, morphemic, or lexical units are arranged according to the rules that are accepted in a given language as norms in order to produce sentences that make sense and are understandable to all the speakers of the language. In his classic 1957 work *Syntactic Structures*, Noam Chomsky puts forward the major principles of syntactic analysis and hence sets the foundations of generative grammar, which imply the generation of innumerable number of sentences by finite number of grammatical rules.

One of the first things to do in analyzing a sentence is to identify groupings within it, the process often referred to as *sentence parsing*. The groupings may range from morphemes to groups of words that

are often called *sense groups*. The groups are called *verb phrase*, *noun phrase*, *adjective phrase*, and so forth. Phrases can in turn be divided into their constituent words. The words can be further subdivided in their constituent morphemes. Therefore, a sentence can be viewed as a hierarchical structure that can be analyzed according to its constituent parts. Furthermore, simple sentences can be joined into more complex structures by different linking devices, such as conjunctions, which are divided into coordinating and subordinating types. For example:

- Tom went to work and met his colleagues (here, *and* acts as a coordinating conjunction).
- Tom went to work, because he had to meet his colleagues (here, *because* acts as a subordinating conjunction).

In modern linguistics, many linguists emphasize the concept of *fuzziness* in syntactic structures. *Fuzziness* implies the production of syntactic structures that are not quite grammatical but are not totally ungrammatical either, meaning that individual speakers produce syntactic structures that do not strictly follow grammatical rules but neither are they incorrect.

Sentence Analysis Through Immediate Constituents and Parsing

One of the most widely used techniques of syntactic analysis is the division or splitting of a sentence into its immediate constituents. The process is usually represented through the construction of a diagram that vividly displays the hierarchical composition of the parts of a sentence and their interrelation with each other. However, the problem with immediate constituent analysis is the fact that the diagram does not inform us about the identity of the sentence elements that they disclose, nor does it reveal the interrelation between the sentences (i.e., active-passive, statement-question). In order to address the issue and make the analysis more informative, it has been suggested to label the parts of a sentence that are disclosed, that is, to indicate the subject, object, and predicate of a sentence.

Another way to approach the analysis is the aforementioned method of *parsing*, or identifying

sense groups within the sentence, such as noun phrase (NP), verb phrase (VP), and so forth, thus making the analysis more informative and thorough.

Cohesion and Coherence

The most important binding devices that hold the syntactic structures are cohesion and coherence. Coherence is achieved through syntactical features such as the use of *deictic*, *anaphoric*, and *cataphoric* elements or a logical tense structure, as well as presuppositions and implications connected to general world knowledge.

Regarding cohesion, M. A. K. Halliday and Ruqaiya Hasan identify the following categories of cohesive devices that create coherence in a text: reference, ellipsis, substitution, cohesion, and conjunction. *Anaphoric* and *cataphoric references* are the significant constituents of building coherent and cohesive connections between and within sentences. *Anaphoric* reference takes place when the reference is made to someone or something that has already been mentioned previously. *Cataphoric* reference is the opposite of anaphoric reference; that is, it refers to someone or something that will be mentioned after the sentence under discussion. More specifically, something is introduced before it is identified. Referents can help us disambiguate a sentence with an indefinite and another operator by anaphoric linkage. Referential anchoring means that the referent of the specific NP is functionally dependent on the referent of another expression. Thus, specificity and familiarity conditions are created that act as binding tools in building the cohesive texts and structures.

Elliptical sentences leave out their constituent parts because the surrounding structures make it possible to make sense of the meaning that is expressed or implied by the sentence.

In the case of *substitution*, a word is not omitted, as in ellipsis, but is substituted for another, more general, word. For example, *which book would you like? I would like the one on the top shelf*, where *one* is used instead of repeating *book*. The same holds for pronouns, which replace the noun. For example, *book* is a noun, and its pronoun could be *it*, as in, *I bought the book because it was interesting*.

Regarding *cohesion*: In linguistics, *grammar* refers to the logical and structural rules that govern the composition of clauses, phrases, and words in any given natural language. The term refers also to the study of such rules, and this field includes morphology and syntax, often complemented by phonetics, phonology, semantics, and pragmatics.

As previously mentioned, *conjunctions* express coordinating and subordinating connections between the parts of complex sentences. A conjunction establishes a relationship between two clauses. The most basic but least cohesive is the conjunction *and*, which mainly expresses coordinating connections. Furthermore, *transition* words are conjunctions that add cohesion to text and include *then*, *however*, *in fact*, and *consequently*. Conjunctions can also be implicit and deduced from correctly interpreting the text.

Subordination and Coordination

Coordination and subordination are ways of making sentences more complex. In coordination, the clauses that are linked are of equal grammatical status, whereas in the case of subordination, also known as *embedding*, one clause functions as part of another—that is, part of the main clause—and is dependent on it. The following examples serve to illustrate the point: *John went to Paris and Nick went to Rome* (the case of *coordination*, i.e., both clauses have equal status); *John went to Paris, because he did not want to go with Nick to Rome* (the case of *subordination*, in which the second clause is subordinate to the main clause).

Mariam Orkodashvili

See also Linguistics, Contemporary; Linguistics, Historical; Semantics (Introduction to Theory); Semantics (Scientific and Empirical); Syntax (Scientific and Empirical)

Further Readings

- Aarts, B. (2007). *Syntactic gradience: The nature of grammatical indeterminacy*. Oxford University Press.
 Aarts, B., Denison, D., Keizer, E., & Popova, G. (Eds.). (2004). *Fuzzy grammar*. Oxford University Press.

- Boeckx, C. (2006). *Linguistic minimalism. Origins, concepts, methods, and aims*. Oxford University Press.
- Chomsky, N. (1957). *Syntactic structures*. Mouton.
- Chomsky, N. (1981). *Lectures on government and binding*. Foris Publications.
- Chomsky, N. (2007). *On language*. The New Press.
- Crystal, D. (1995). *The Cambridge encyclopedia of the English language*. Cambridge University Press.
- Crystal, D. (1997). *The Cambridge encyclopedia of language*. Cambridge University Press.
- Haegeman, L. (2005). *The syntax of negation*. Cambridge University Press.
- Halliday, M. A. K., & Hasan, R. (1976). *Cohesion in English*. Routledge.
- Heusinger, K. von. (2002). Specificity and definiteness in sentence and discourse structure. *Journal of Semantics*, 19, 245–274.
- Karttunen, L. (1976). Discourse referents. In J. McCawley (Ed.), *Syntax and semantics: Vol. 7. Notes from the linguistic underground*. Academic Press.
- Widdowson, H. G. (1996). *Linguistics*. Oxford University Press.
- Woods, A., Fletcher, P., & Hughes, A. (1986). *Statistics in language studies*. Cambridge University Press.

SYNTAX (SCIENTIFIC AND EMPIRICAL)

Language is a uniquely human capacity. Within this capacity, organizing words into logically coherent and comprehensible structures—sentences—is the most complex human faculty. These sentences, or syntactic structures, present the core of syntactic analysis, more precisely, of syntax. The major concern of syntax is the study of syntactic structures in all their varieties and complexities. Syntax may lie within empirical and scientific frameworks.

Empirical Framework of Syntax

The empirical framework of syntax is concerned with the practically observable processes of arranging words into sentences, identifying parts of sentences, and parsing sentences into sense groups and other readily analyzable parts. One of the major realms of empirical syntax is syntactic typology.

Syntactic Typology

The order of subjects, verbs, and objects in sentences determines the typological classification of languages according to syntactic structures. Apart from variation in the basic order between subjects, verbs, and objects, languages can also reveal other types of variation. As an instance, some languages show more word-order possibilities than others. This may be attributed to morphological properties; that is, how much case morphology exists in noun phrases in the language. Generally speaking, languages with well-developed case morphology have a freer word order than languages with fewer case categories (e.g., Caucasian/Kartvelian/languages with free word order, as opposed to Indo-European languages with fixed word order in a sentence).

Intuitively, it is easy to understand the reasons behind this fact. Case morphology readily indicates the function of a noun phrase, whereas in the absence of case morphology, the language has to look for alternative ways of indicating which constituent has subject function, which is the direct object, and which constituent is the indirect object. Implementing a strict word order carries out this duty of identifying the functions of the constituent parts of a sentence.

Hence, we can classify the languages according to syntactic typology; in other words, through identifying the types of syntactic structures that are characteristic of a certain language we can attribute it to a specific class of languages that have similar syntactic structures. Taking a step further, syntactic typology allows for identifying universals across languages that in its turn facilitates cross-language comparative analysis.

Clause Grammar and Structure

Another significant field of empirical syntax is that of clause grammar and structure. It is well known that almost all languages contain complex syntactic structures alongside simple sentences that possess straightforward and basic syntactic structures. Therefore, syntax often concerns itself with clause grammar rather than sentence grammar. This is primarily because complex sentences are made up of clauses that are easily subject to syntactic analysis. A clause is a unit structured

around a verb phrase. Together with the verb phrase, there are clause elements, which are realized by phrases or by embedded clauses. The core of the clause can be divided into two main parts: the *subject* and what is traditionally called the *predicate*. These usually correspond to a nominal part and a part with a verbal nucleus that semantically is equivalent to a *topic* and a *comment*. Together, these parts express a proposition. The predicate in its turn can be broken down into a verb phrase and several complements (objects, predicatives, adverbials). Thus, the main constituent elements of a clause are subject, verb phrase, subject predicative, direct object, indirect object, prepositional object, object predicative, and adverbial. Each element plays a role in the syntactic analysis of a clause and in deriving its meaning and its relations to surrounding structures.

Subordinate clauses may have various syntactic functions, such as nominal clauses, exclamative clauses, relative clauses, adverbial clauses, and so forth. Identification of the structural and functional features of clauses and the analysis of complex sentences facilitate the formulation of principles that help us trace the development and evolution of languages.

Syntactic Gradience and Fuzzy Syntactic Structures

In modern empirical linguistics, a growing number of linguists emphasize the concept of *fuzziness* in syntactic structures. *Fuzziness* implies the production of syntactic structures that are not quite grammatical but are not totally ungrammatical either; individual speakers produce syntactic structures that do not strictly follow the rules of grammar but neither are they incorrect. Hence, we get the condition of *syntactic gradience*, whereby the boundaries between grammatical and nongrammatical structures have become blurred. Therefore, more scientists hold the positivist stance toward syntactic analysis of structures by merely *describing* the syntactic idiosyncrasies of languages and producing the findings on usage frequencies and conditions of certain constructs rather than standardizing the norms. The shifty nature of languages offers more possibilities for positivist studies rather than normative standardization of grammatical forms,

be they on the level of phonology, morphology, syntax, or semantics.

Conditional Syntactic Structures

Conditional sentences offer several interesting insights into empirical syntax, namely, the syntactic construction of special types of sentences. *If*-clauses standardly have the form of declarative sentences and are introduced into the scope of *if* without changes. However, sometimes the clauses of a conditional construction are elliptical, meaning that they are missing elements that can be surmised from the context. In most cases, elliptical *if*-clauses (and sometimes even main clauses) are interpreted exactly like their full forms and seem to be reduced merely to avoid repetition. For instance, the following constructs illustrate the point: *I'll go, but not if John goes. If not, why not? If not now, then when?* However, there is a class of *if*-clauses that require other interpretations, since they are not conditionals per se, but rather illustrate the structures of concessive clauses. For instance, the following clauses are concessive and not conditional: *The Queen of England is happy, if not ecstatic. He spoke ungraciously, if not rudely. She ate many of the cakes, if not all. The salary was good, if not up to her expectations.*

Further interesting point about the construction of conditional sentences is the case of *inversion* which is used to make the sentence more emphatic. The following could serve as examples: *Should you see John, tell him I am here. Were you better prepared, you would sit for an exam now. Had you warned us before, we would have arrived earlier.* The placement of verbs at the beginning of a sentence produces more emphatic constructs that acquire additional emotional connotations through specific syntactic structures.

Scientific Framework of Syntax

Syntactic structures, or sentences, help humans to express thoughts, abstract and concrete notions and phenomena, desires, intentions, and goals. To put it differently, sentences reflect human cognitive processes, such as thinking, problem-solving, decision-making, and perhaps most importantly, creativity.

Therefore, syntax lies within the framework of science, namely that of cognitive science. It applies the scientific methods of gathering and observing data; of making generalizations based on the observed data; and thereafter, of developing hypotheses. All these steps are characteristics of scientific fields such as syntax, among many others.

Through innumerable sentences can be expressed an innumerable number of thoughts, ideas, and opinions that reflect human cognition. Observing these verbally structured cognitive processes enables scientists to make generalizations about human cognition, thinking, learning, memory, concentration, and the logical, coherent expression of one's ideas.

These data are gathered through observation of a vast range of syntactic structures uttered or written by humans in a huge number of instances. Thereafter, the method of grammaticality judgment is applied to distinguish coherent and well-structured sentences from those that are incoherent and ambiguous.

Through the analysis of syntactic structures, through their simplicities or complexities, the development and evolution of human cognition and human language can be traced over extended periods of time.

Therefore, syntax, as a scientific method, vividly reflects, analyzes, and generalizes human and linguistic evolutionary stages.

Transformational and Generative Grammar

The generation and transformation of syntactic structures into their numerous modifications was proposed by the linguist Noam Chomsky, who attributed the capacity of the production of innumerable sentences through a finite set of grammatical rules innate in human nature to acquire and use language. In his landmark study on *Syntactic Structures*, Chomsky suggests a model combining *phrase structure* and *grammatical transformations*, which has been demonstrably effective for syntactic analysis. The most significant activity in this respect is tracing the chains of sentence transformations and following the entire process of generating various syntactic structures by applying grammatical rules. Another significant Chomskyan contribution to linguistics is the concept of deep structure, as opposed

to surface structure, which enables more sophisticated ways of syntactic analysis.

Government and Binding

Within transformational and generative grammar, there exist several theories essential for syntactic analysis. Central to syntactic analysis are the theories of government and binding. *Binding theory* poses locality conditions on certain processes and related items, whereas the central notion of government theory is the relation between the head of a construction and categories dependent on it. *Theta* (θ) *theory* is concerned with the assignment of thematic roles such as agent-of-action (i.e., θ -roles). Binding theory is concerned with relations of anaphors, pronouns, names, and variables to possible antecedents. *Case theory* deals with assignment of abstract case and its morphological realization. *Control theory* determines the potential for reference of the abstract pronominal element. These subsystems are closely related in a variety of ways. For instance, binding and case theory can be developed within the framework of *government theory*. Furthermore, case and θ theory are closely interconnected. Through the interaction of these systems, many properties of particular languages can be accounted for. A great deal can be learned about Chomsky's notion of *universal grammar* (i.e., that structural rules of language are innate in humans) from the study of a language, if such study achieves sufficient depth to put forth principles that have explanatory force but are underdetermined by evidence available to the language learner. The study of closely related languages that differ in some clustering of properties is particularly valuable for identifying and clarifying parameters of universal grammar that permit variations in the proposed principles.

Syntax of Negation

Negative constructs present an especially interesting case for syntactic analysis, which mainly falls within government-and-binding framework. Negative structures and their development over the course of centuries bear significance for word-order alterations in sentences. The exclusion of

double negation in modern English is one such example of tracing developmental stages in language evolution. Negative constructs also play significant role in the comparative cross-language studies.

Statistical Analysis

In modern linguistics, statistical analysis has been widely used for measuring in quantitative terms the acceptability of various syntactic structures. Obtaining data for syntactic analysis has presented a special challenge to researchers. Sampling decisions have been particularly tough to make, since syntactic structures present a special case where the data need to be extracted from representatives of native speakers in order to evaluate the validity of the *grammaticality* of a construct. In syntax research, an interval scale of grammaticality is commonly used. Sentences are rated as grammatical, questionable, or ungrammatical. However, it has been observed in the study of sensory impressions that raters are more consistent with an open-ended rating scale than they are with category rating scales. So, in recent years, methods from the study of psychophysics (subjective impressions of physical properties of stimuli) have been adapted in the study of sentence acceptability. Hence, researchers are evaluating the *degrees of relative acceptability* of a sentence rather than grammaticality or ungrammaticality at its extreme forms. The main concern of researchers at this point is measuring the degree of word-order violations in various syntactic structures and accordingly judging the degrees of their acceptability. Statistical analysis is widely applied both in empirical and scientific frameworks of syntax.

Mariam Orkodashvili

See also Linguistics, Contemporary; Linguistics, Historical; Semantics (Introduction to Theory); Syntax (Introduction to Theory).

Further Readings

Aarts, B. (2007). *Syntactic gradience. The nature of grammatical indeterminacy*. Oxford University Press.

- Aarts, B., Denison, D., Keizer, E., & Popova, G. (Eds.). (2004). *Fuzzy grammar*. Oxford University Press.
- Biber, D., Johansson, S., Leech, G., Conrad, S., & Finegan, E. (1999). *Longman grammar of spoken and written English*. Longman.
- Carnie, A. (2013). *Syntax. A generative introduction* (3rd ed.). Wiley-Blackwell.
- Chomsky, N. (1957). *Syntactic structures*. Mouton.
- Chomsky, N. (1981). *Lectures on government and binding*. Foris Publications.
- Chomsky, N. (2007). *On language*. The New Press.
- Crystal, D. (1995). *The Cambridge encyclopedia of the English language*. Cambridge University Press.
- Crystal, D. (1997). *The Cambridge encyclopedia of language*. Cambridge University Press.
- Dancygier, B. (2006). *Conditionals and prediction. Time, knowledge and causation in conditional constructions*. Cambridge University Press.
- Haegeman, L. (2005). *The syntax of negation*. Cambridge University Press.
- Halliday, M. A. K., & Hasan, R. (1976). *Cohesion in English*. Routledge.
- Heusinger, K. von. (2002). Specificity and definiteness in sentence and discourse structure. *Journal of Semantics*, 19, 245–274.
- Johnson, K. (2008). *Quantitative methods in linguistics*. Blackwell.
- Karttunen, L. (1976). Discourse referents. In J. McCawley (Ed.), *Syntax and semantics 7: Notes from the linguistic underground*. Academic Press.
- Widdowson, H. G. (1996). *Linguistics*. Oxford University Press.
- Woods, A., Fletcher, P., & Hughes, A. (1986). *Statistics in language studies*. Cambridge University Press.

SYSTEMS SCIENCE

Systems science is the ordered arrangement of knowledge whose domain of inquiry consists of *thinking*, *conceptualizing*, and *modeling* the behaviors of the real world in a systematic way. This entry discusses the fundamentals of systems science, including (a) the ways in which systems thinking represents the main characteristics of systems, (b) how system conceptualization has defined the basic concepts of systems, and (c) the role of systems modeling in providing the tools

that are used to create and test systematic experiments to simulate real-world behaviors.

Systems Thinking

Systems thinking is a perspective that considers events and structures as parts of a whole system that carries out a certain function, and which attempts ascribe meaning to systems based on their parts and on the interaction of these parts.

Construction of systems science begins with systems thinking and goes back to the early 1900s. In the 1920s, Ludwig von Bertalanffy developed the concept of *organismic biology* using a somewhat Darwinian approach. His views, however, did not attract significant attention from the scientific community. In 1937, Bertalanffy presented a report titled “The General Systems Theory (GST).” Although this report also failed to attract much attention at the time, it nevertheless has come to be recognized as one of the most important studies in the field of science. It was not until 1950 that Bertalanffy was able to publish his first article, in which he clarified and elaborated his thoughts regarding the GST.

During the 1954 meeting of the American Association for the Advancement of Science, a new society was founded by the biologists Bertalanffy, Kenneth Boulding, Ralph Gerard, James G. Miller, and Anatol Rapaport. This society was initially named the Society for the Advancement of General Systems Theory and later renamed as the Society for General Systems Research.

GST has its roots in the concept of *organismic biology* developed by Bertalanffy: In contrast to physical forces such as gravity and electricity, life cannot exist outside of individual entities known as *organisms*. Each organism is a system, and thus a dynamic order with its own interactions, parts, and processes.

Both organic and mechanical systems can be the subject of systems thinking. For example, the systems thinking used to describe the relationship between the energy spent and the distance covered by a lizard when it is jumping can similarly be used to describe the relationship between the fuel consumed by a car and its movement.

What matters in this context is not so much the similarity of the “objects” or “cases” being

considered but rather the similarity of processes. For example, the point that needs to be considered is the similarity between the exponential growth processes for money and the exponential growth processes of human populations. Therefore, in case money is placed on interest at a yearly rate of 7%, and in case a population similarly grows at a rate of 7% every year, both the sum placed on interest and the population in question will double in 10 years. When determining the behavior of a system, the structure of the system (the order and nature of the processes) is generally as important as its components.

In addition, both natural and artificial systems can be evaluated within the scope of systems thinking. For example, just as it is possible to explain the functioning of the human body by using systems thinking, so can the organization and functioning of a business also be explained by using systems thinking.

According to perspectives based on systems thinking, the structures that constitute living organisms are the result of interactions between parts that possess their own inherent characteristics. Thus, by defining these parts and their interactions, it becomes possible to understand the living organism. A component, or part, is an entity that has its own characteristics and which can become associated with other components/parts. When different parts come together or are brought together to perform a certain function, this results in the formation of a whole entity known as the *system*.

Systems also bear and reflect the characteristics of their parts. Depending on their function, they will also have their own inherent characteristics. Different systems can also be assembled together to form different and even more complex systems.

From this perspective, many of the entities encountered in our lives are systems. Entities that are not systems are referred to as *masses*. The main difference between masses and systems is the way in which they perform their functions. A mass is limited to its constituting parts. As such, the interactions between these parts do not create a special function. For example, masses of sand, water, or cement are formed by grains of sand, droplets of water, and grains of cement. However,

because these constituents do not have a particular function, the characteristics of masses will not change regardless of how many parts the initial masses are divided into. If one divides a sand pile into two parts, and then into four parts, one will obtain masses which, despite their smaller sizes, will have the same characteristics as the initial mass. On the other hand, if sand, water, and cement are brought together in correct proportions within a building system, the relationship between these parts will create not only a new function but also a new system. In such a case, dividing the resulting system into smaller parts will cause the system to lose its ability to fulfill its function. For this reason, *a system represents a whole that is more than the sum of its parts.*

The relationships between the parts that constitute a system are causal with regard to events and symbiotic with regard to structure. A causal relationship refers to the cause-and-effect relationship between the interacting parts. A causal relationship applies when, within the context of an event, a cause in one part of the system creates an effect at another part of the system. A symbiotic relationship refers to those cases when there is a mutual interdependence in the relationships of the different parts constituting the system. With regard to systems science, the basic principles that determine the characteristic of the relationships between the different parts are as follows:

1. The relationships are of a dynamic character: Relationships constantly evolve to the state in which the system's function can be best fulfilled.
2. The relationship between the different parts can involve the flow of materials, information, or energy. The same rules are applicable under different circumstances and conditions.
3. The relationships are sensitive to environmental changes.
4. Together, the system's relationships form a dynamic relationship.
5. The system relationships are based on cause-and-effect relationships.
6. Relationships indicate how one part can affect another part within the same system. For example, increasing urbanization leads to a decrease in agriculture. Excitement causes blood

pressure to rise. In system relationships, the first variable influences the second variable.

In such a relationship structure, it is possible to describe two different relationship groups. The first relationship group refers to the group of relationships between the parts that constitute the system. For systems, the first group relationships represent input–output relationships. All the values that systems require to fulfill their own functions are considered as inputs. On the other hand, the products obtained following the processes and functions of these systems represent the outputs.

Systems Concepts

Systems concepts are definitions of specific functions or parts of the systems. These definitions help us to understand the behaviors of the systems.

In case a system does not have an input–output relationship with its environment, this system is referred to as a *closed system*. Systems that have input–output relationships with their environment, on the other hand, are known as *open systems*. Within the scope of these relationships, the structure that exists between the inputs and outputs of the system will indicate whether the first group relationships are linear. To determine whether the relationships within this group are linear, two tests are performed:

- **Homogeneity test:** In case a change in one of the inputs of the system causes an equal change in the system's outputs, then the first group relationships of the system are considered to be *linear*.
- **Additivity test:** In case the sum of the effects of the two inputs is equal to the sum of the effects caused by these inputs through the system, the system is considered to exhibit a linear relationship. However, in case the sum of the effects of the two inputs is different from the sum of the effects caused by these inputs through the system, the system is considered to exhibit a *nonlinear* relationship.

Systems with fully linear relationships are known as *linear systems*, while systems without linear relationships are as *nonlinear systems*.

The other relationships of a system include the relationships between its own parts. In case the changes in the relationships between a system's parts result in predictable and definable changes in the system's behavior, the system is considered to display simple relationships. Systems with such relationships are known as *simple systems*. In case the changes in the relationships between a system's parts result in unpredictable and indefinable changes in the system's behavior, the system is considered to display complex relationships. Systems with such relationships are known as *complex systems*.

Ross Ashby's *law of requisite variety* demonstrated that a complex system can only be controlled by another system that possesses a mechanism of equal complexity. According to Ashby's approach, "only variety absorbs variety." More complex systems are more sensitive to their environments. The minimum variety necessary to perceive a complex environment is also determined to be a system's complexity. Social organizations are more complex than physical organizations/systems.

A system undergoes an adaptation process that depends on the structure and nature of the environment in which it is found. This structure is dynamic. As a result, part of the change observed in the system will stem from its efforts to adapt to the new environment, while the remaining changes will stem from the attitude and behavior of other systems. This interactive aspect of the environment is reminiscent of the processes that biological systems undergo. Within this structure, the survival efforts of systems are limited to their possible behaviors. The choices of a system within changing structures will determine its likelihood of surviving. These types of environments are called *fitness landscapes* and are very well known from the study of biological systems. For this reason, the first studies and definitions for fitness landscapes have been defined and developed many years ago (first by Herbert Spencer in 1864).

The fitness landscape is composed of the probabilities for different decisions that the organization might make during a change process. Stuart Kauffman previously demonstrated that this landscape could be defined by using the number

of these probabilities (N) and the connectivity of each choice (K). Kauffman's work thus allowed the fitness landscape model, which is also known as the *NK model*, to acquire a mathematical expression. Each choice probability has a corresponding fitness value. This value (which is the mean of all probabilities) describes the area in which a choice is located within the system and with respect to other choices (i.e., its distance and relationship to other choices). Kauffman defined the fitness landscape mainly for biological systems.

Self-organization increases when systems interact with their environments at the edge of chaos. The flexibility provided by self-organization renders it easier for successful organizations to constantly adapt to the changes in their environment. Three important mechanisms that allow systems to self-organize are identification, information, and relationships. A system formed by self-organization acquires certain characteristics and behaviors, as well as certain spontaneous properties.

In case a system is a closed system, its function (commonly understood as its reason for existence) will not be dependent or affected by its interactions with the environment. In other words, the system's functions will not be dependent on inputs and outputs. In such a system, the outputs will gradually decrease if the inputs remain the same. This, in turn, will cause the system's energy to constantly decrease. For this reason, closed systems eventually lose the qualities that allowed their constituent parts to remain together, which leads them to fall apart. This tendency of closed systems to fall apart is defined as *entropy*.

Entropy reflects the fact that the functions of systems are not associated with their environmental interactions and that they consequently experience a constant loss of energy that ultimately leads them to fall apart. Thus, to strengthen the bonds between the parts of a system, it is necessary to strengthen the relationship between its function and its inputs–outputs. As such, it is necessary to increase the inputs or the outputs relating to the function. The flow of energy that leads to an increase in the system's inputs or outputs is called *negative entropy*.

To be able to perform their functions in a sustainable manner, systems attempt to increase their

input–output relationships. This allows them to continue their existence—or to continue living. In this context, *living systems* are systems that can constantly develop negative entropy.

The efforts demonstrated by systems in order to increase their negative entropy are known as *adaptation*. Systems that can undergo change in order to develop their input and output relationships are called *adaptive systems*.

The relationships and interactions of systems in both groups are bidirectional. The fact that these relationships are bidirectional means that both the relationships between the system and its parts and the relationships between the system and its environment will have an effect on the system. These effects can reshape the behavior of the system with respect to its activities. This type of interaction is called *feeding*.

Feeding that stems from the relationship between the environment and the system is called *feedforward*. Feedforward involves an evaluation process that progresses from system inputs to system outputs (from backward to forward). Before beginning its functions, a system first attempts to obtain inputs from the environment that are suitable for its functions. This allows the system to shape itself according to future conditions when developing its functions.

Feeding that stems from the relationships between the system and its parts is called *feedback*. Feedback involves an evaluation of whether the system acts according to its functions through assessments that progress from the system's outputs to the system's inputs (from forward to backward).

In case a system function deviates from its previously determined goals, it becomes necessary to perform a correction in the inverse ratio and direction of the deviation. In case this correction increases/improves the function of the system, the feedback is described as *positive feedback*. In case this correction decreases/reduces the function of the system, the feedback is described as *negative feedback*.

Feedback allows systems to reach a state of equilibrium. An important feature of systems that can reach equilibrium is their ability to react actively to change. Most living systems continue to function and to use energy even when they are not required to display any reactions. Mechanical systems can be entirely deactivated for a certain

period of time and then be reactivated. The same type of behavior is also observed in certain simple life forms. Certain plant roots can wait for years until suitable conditions for growth become available, while certain microorganisms can be frozen and then reanimated. However, most organisms need to remain active in order to stay alive. In other words, they will inevitably die if they become inactive.

As a general rule, a more complex system will also have a higher level of internal activity. This increases the likelihood of the system to trigger change in its environment. Complex systems spend more energy to process information and to ensure their continuity. This same rule is also applicable for social systems. For example, while a club may be easily closed and then reopened, doing so for a company is far more difficult. A society, on the other hand, needs to be constantly active and to ensure its own continuity.

The behavior of systems is largely determined by their feedback reactions. This phenomenon is called the *reaction time*. The functioning of a simple thermostat/oven involves a long reaction time, since air conveys heat from the oven to the thermostat at a relatively slow rate. Social systems also have generally long reaction times. If the reaction time is too slow, external changes will take place at a comparatively faster rate, such that the system will be unable to respond. In addition to this, the reaction time is generally the same as the basic tempo, or the “shot” of the system (the minimum interval between the oscillations of the feedback process). By watching these oscillations and estimating the time interval, it is possible to determine the minimum time interval that a certain system may overcome.

Forming the decisions of a system also requires its identity to be formed/created. The property of a system that allows sufficient free movement for the formation of an identity is referred to as its *viability*.

Systems Models

Systems models are tools that are representations of the real world in different perspectives. There are different types of systems models based on their perspectives. Some of them are explained briefly as follows.

Complex Adaptive Systems (CAS) Model

H. John Holland argued that integrating systemic complexity and environmental complexity with one another could only be achieved on the basis of the inherent abilities of the system. Foremost among these system abilities is the adaptation skill. For this reason, such systems are referred to as CAS. The general structure of the CAS is framed according to seven basic laws. Four of these laws describe the characteristics of the CAS. These are the laws on aggregation, nonlinearity, flows, and diversity. The other three laws describe the mechanisms of the system, which are tagging, internal models/schemas, and building blocks. CASes can be described as the fundamental elements of the concept of complexity. CASes are self-organizing systems. Self-organizing systems tend to organize themselves by using energy, materials, and feedback from their internal and external environments. Similar to traffic, crowdedness, and organization, self-organization is also a characteristic of a multicomponent system. Self-administering teams within organizations represent examples of self-organization.

Social Systems Model

Luhmann has contributed to systems science in the area of sociology by constructing the social systems model. Luhmann's views were largely based on the structural functional systems theory, whose foundations were established by Talcott Parsons. Parsons associates the existence of a system to its function. For the continuity of the function, systems establish their own optimal balances. To ensure balance, it is important to closely follow the system's relationships with the environment and to ensure that the parts of the system remain together as a whole. It is for this reason that institutions are necessary. With this type of structure, all events that occur within society are the result of interactions between the actors. The action that forms the basis of this interaction harbors four basic elements/aspects: adaptation, goal access, integration, and implicit pattern formation.

Luhmann systems are divided into two basic groups, which are the autopoietic and the

allopoietic groups. *Autopoietic group systems* are systems that undergo structural changes of their own when establishing a dynamic balance. These include social (interaction, organization, and social) systems as well as psychological systems. *Allopoietic systems*, on the other hand, refer to machines and organisms.

System Dynamics Model

Toward the end of the 1950s, Jay W. Forrester developed a new theory based on concepts relating to feedback, work evaluation, organizational and social relationships, and the structural changes associated with these relationships. The theory involved the modeling of mathematical changes in structural movements.

The rationale of the system dynamics model is based on the view that the movements or changes triggered in real structures will actually consist of small-scale movements, while complex situations mainly stem from relationships and interactions between the elements of the system. The interaction of all effective system elements on any structure will lead to complexity. It is believed that this complexity can be understood through the analysis of feedback structures that are inherent to all relationships. For this reason, feedback structures play a key role for understanding complexity.

In this context, the graphic illustration and description of the system dynamics model indicates that this model is very similar to GST. In addition to being easy to use, the system dynamics model allows the modeling of nonlinear relationships within an organization. It thus represents a tool that can be easily used and understood by those in decision-making positions. Although models developed using system dynamics involve numerous elements and relationships, they do not take much time to prepare and are quite descriptive.

The time-dependent change observed in the elements or variables that constitute a system represent the behavior of a system. The behavior of a system is intrinsic, dependent on the baseline, and nonlinear. Systems demonstrate their behavior under the influence of their own structure and elements. This behavior is displayed within the limits and effects of the system's structure. Based on this consideration, it is possible to show that a system's

structure also determines the behavior of the system. Thus, system dynamics identifies the relationship between a system's behavior and the structure that is at the source of this behavior. In other words, system dynamics evaluates how the structure of a physical, biological, or literary system shapes the displayed behavior of the system. For example, by identifying the structure of an ecological system, system dynamics allows the time-dependent behavior of this system to be effectively analyzed.

Viable Systems Model

Viable systems modeling was constructed by Stafford Beer. It represents a cybernetic approach to systems science. According to Beer, freedom does not define a state that is distant or dissociated from the other entities in the environment. For example, what defines a university is the education it provides and the research it performs. Without these functions, an institution would not be considered a university. As such, all of the units, faculties, departments, centers, and so on, that provide education, conduct research, and bring life to the university through their products/outputs can be considered living systems. The other structures within the university such as dormitories, dining halls, and councils and committees, on the other hand, are not living systems. For a system to be considered a living system, it must be able to interact in autonomously with its environment. Autonomy allows a system to make decisions based on its own abilities/capacities. Viable systems have mission statements, as well as an accountabilities within the scope of the functions they carry out in accordance with the policies and strategies of the larger entity of which they are part. The continued viability of systems is achieved by two main mechanisms: (1) cohesion and (2) adaptation. The cohesion mechanism aims to ensure that the goals and interests of the independent units within the main system are in agreement with the policies, strategies, goals, and interests of the main system and that the competitive strength of these units are reinforced. The adaptation mechanism, on the other hand, aims to ensure that the system is able to adapt in a timely manner to any changes that might occur in the external environment. In addition to adaptation,

learning, development, and evolution are considered among the basic dynamics of the concept of viability. It would not be possible to achieve evolution without first achieving adaptation; to achieve adaptation without first achieving learning, and to achieve learning without first achieving equilibrium.

Causal Relations Model

Causal relations models are system models that represent systems behavior with causal relations. Causal relations do not refer to a ranking but to influences on system functions. The relationships within the system are not linear but circular. The circular definitions of system behavior is called a "structure." Different real-world behaviors can be modeled with the same structure. In this perspective, it is accepted that structures define the behaviors.

A systems structure can be constructed according to two different kinds of loops. The first kind of loops are *reinforcing loops*. Such loops generate either a continuous increase or decrease in systems behavior. For example, the number of customers is increased by increasing the investment in products or services. Increase in number of customers leads to increase in revenues, which leads to increase in investments. The other type of loops are *balancing loops*. In balancing loops, at least one (or in odd numbers of) relation is a negative relation. For example, if the positive perception of a service quality is increased, more people are attracted, thereby creating a queue. While the number of the people waiting in the queue is increasing, the waiting time is increased. Increase in waiting time leads to decrease in the perceived service quality. Decrease in perception about service quality leads to decrease in the number of the customers attracted. This loop models a systems behavior that is opposite to the initial behavior.

In causal relations models, it is accepted that delays in relations and mental models of the actors can create the changes in systems behaviors. Mental models that cause changes in relations are represented by clouds. Delays are represented by two parallel lines.

Yavuz Ercil

See also Complex Systems; Developmental Systems Theory; Perturbation Theory

Further Readings

- Ackoff, R. (1974). *Redesigning the future: A systems approach to societal problems*. John Wiley & Sons.
- Ashby, W. R. (1956). *An introduction to cybernetics*. Chapman and Hall.
- Ashby, W. R. (1960). *Design for a brain: The origin of adaptive behavior*. Chapman and Hall.
- Ashby, W. R. (1962). Principles of the self-organizing system. In H. von Foerster & G. Zopf (Eds.), *Principles of self-organization* (pp. 255–278). Pergamon Press.
- Beer, S. (1985). *Diagnosing the system for organizations*. John Wiley & Sons.
- Churchman, C. W. (1968). *The systems approach*. Dell.
- Emery, F. E. (Ed.). (1969). *Systems thinking: Selected readings* (Vols. 1 and 2). Penguin.
- Flood, R. L. (1989). Six scenarios for the future of systems “problem solving.” *Systems Practice*, 2, 75–99.
- Flood, R. L., & Jackson, M. C. (1996). *Critical systems thinking: Directed readings*. John Wiley & Sons.
- Forrester, J. W. (1961). *Industrial dynamics*. MIT Press.
- Forrester, J. W. (1971). Counterintuitive behavior of social systems. *Technology Review*, 73(3), 53–68.
- Forrester, J. W. (1994). System dynamics, systems thinking, and soft or. *System Dynamics Review*, 10(2–3), 245–256.
- Forrester, J. W. (1975). *Collected papers of Jay W. Forrester*. Productivity Press.
- Gell-Mann, M. (1995). *The quark and the jaguar*. W. H. Freeman.
- Gleick, J. (1987). *Chaos: Making a new science*. Viking.
- Holland, H. J. (1995). *Hidden order: How adaptation builds complexity*. Perseus Books.
- Kauffman, S. (1993). *The origins of order: Self organization and selection in evolution*. Oxford University Press.
- Kim, D. (1994). *Systems thinking tools: A user's reference guide*. Pegasus Communications.
- Luhmann, N. (1982). *The differentiation of society* (S. Holmes & C. Larmore, Trans.). Columbia University Press (Original work published 1976 in German as: *Soziologische Aufklärung*, Westdeutscher Verlag, Opladen).
- Meadows, L. D., Behrens, W. W., III., Meadows, D. H., Naill, R. F., Randers, J., & Zahn, E. K. O. (1974). *Dynamics of growth in a finite world*. Wright-Allen Press.
- Midgley, G., (Ed.). (2003). *Systems thinking: Volume I & II. General systems theory, cybernetics and complexity*. Sage Publications.
- Peters, T. J. (1987). *Thriving on chaos*. Knopf.
- Prigogine, R. I., & Stengers, I. (1984). *Order out of chaos*. Bantam Books.
- Richmond, B. (1994). Systems thinking system dynamics—Let's just get on with it. *System Dynamics Review*, 10(2–3), 135–157.
- Senge, M. P. (1990). *The fifth discipline: The art and practice of the learning organization*. Doubleday/Currency.
- Senge, M. P., Kleiner, A., Roberts, C., Ross, R., Roth, G., Smith, B., & Guman, E. C. (1999). *The dance of change: The challenges to sustaining momentum in learning organizations*. Doubleday Currency.
- Shannon, C. (1964). *The mathematical theory of communication*. University of Illinois Press.
- Stacey, D. R. (1992). *Managing the unknowable: Strategic boundaries between order and chaos*. Jossey-Bass.
- Sterman, D. J. (2000). *Business dynamics: Systems thinking and modeling for a complex world*. McGraw-Hill Higher Education.
- Sterman, D. J. (2001). System dynamics modeling: Tools for learning in a complex world. *California Management Review*, 43(4), 8–25.
- Sterman, D. J. (2002). All models are wrong: Reflections on becoming a systems scientist. *System Dynamics Review*, 18(4), 501–531.
- von Bertalanffy, L. (1962). General system theory—A critical review. *General Systems*, 7, 1–20.
- von Bertalanffy, L. (1968). *General system theory: Foundations, development, applications*. G. Braziller.
- Waldrop, M. M. (1992). *Complexity: The emerging science at the edge of order and chaos*. Simon & Schuster.
- Weiner, N. (1961). *Cybernetics: Or, control and communication in the animal machine*. MIT Press.
- Weiner, N. (1998). *The human use of humans: Cybernetics and society*. Da Capo Press.



TAXONOMY

Taxonomy is the science of classification and may refer to either the process or the product of a classification system. Most widely used in the biological sciences, taxonomy can refer to any controlled, hierarchical vocabulary attempting to achieve a standardized classification or naming system. Taxonomy, as a conceptual tool, has allowed widespread scientific and technical vocabulary to transcend regional nomenclature and folk knowledges by standardizing the use and relationships of terms. It also can refer to a methodological approach to classification, specifically in botany and zoology. The most widely recognized taxonomic system in the sciences, within botany and zoology, is based on the work of the Swedish botanist Carl von Linné (1707–1778), commonly known as Carolus Linnaeus. He is generally considered the father of modern taxonomy, though many before him had sought to form a standard classification system.

Taxonomies have been studied as tools for conceptual categorization of the world. Folk taxonomies have been of particular interest in psychology and anthropology as conceptual knowledge of things is expressed in cultural systems. While the study of folk knowledge and folk taxonomies are of particular interest in the social sciences about how people act, believe, or think through folk classifications, taxonomy in the biological sciences is more of an active process of classification. Thus, taxonomy can also refer to a conceptual tool of

the mind, a cultural tool of an ethnolinguistic group, or on the level of an internationally standardization project.

Taxonomy as Conceptual Categories

Taxonomies, in the broadest sense, are systems of conceptual categories. They are mental and cultural tools used to help humans understand the world around them. All groups of people have some system of classification of things, and the way things are classified can indicate much about beliefs and values. Anthropologists and linguists, among others in the social sciences, may collect the terms and relationships from a particular ethnolinguistic group. Although the categories may not be scientifically defined, they present a particular view of the world from the perspective of the people within the group. These taxonomies, or classification systems, help to identify the names, categories, and relationships of terms in a particular language. For example, the difference between astrology and astronomy entails a difference in the classification of the celestial bodies. Astrology is a folk system observing the stars and planets based on their spatial relationships from the viewpoint of the observer (i.e., constellations). Although these categories do not match the modern scientific categories within astronomy, historically they have provided some useful information by way of mapping out the relationships of visible celestial bodies through a different, although internally consistent, taxonomy.

Taxonomy as a Controlled Vocabulary

Natural language lacks the specificity for large-scale technical and scientific work. Ambiguity is introduced through synonyms, homophones, and colloquialisms, thus providing a potential threat to specific and technical reference. A common example in botany is that of the European Mountain Ash. Although the common name would seem to indicate its relationship to the true ash tree, they are actually not so similar. The scientific name for the Mountain Ash is *Sorbus aucuparia*, from the genus *Sorbus*, among the rose family. A true ash tree is from the genus *Fraxinus*, from among the olive or lilac family. Although the common name bears striking similarity, the plants are actually distant in the taxonomic hierarchy. Thus, to be specific in scientific classification, a taxonomy in which the names of plants do not bear overlap between categories is preferred.

A taxonomy is an attempt to create a standard classification system for a particular domain in which each term, or *taxon* (plural: *taxa*), would be an individual, specific term hierarchically related to other terms. Taxonomies may be intended to standardize classifications and nomenclature within institutions, throughout a particular industry or field, or even to create the ability for global scientific collaboration. Sometimes, a governing body may be instituted to control the vocabulary to make sure it is standardized, though often the taxonomy is set at a particular point in time and is extended or built upon more organically.

As a standardized vocabulary, a taxonomy is a logical device not always intending to be an empirically defined categorization of the world as such. The limits of language, and the specific use of words-for-things, prevent taxonomies from being cultural maps of the world and rather should be seen as logical categorization processes. Similar to a thesaurus, a taxonomy seeks to control for ambiguity, eliminate synonyms, and to display hierarchical relationships. However, unlike a thesaurus, the taxonomy does not draw lateral or associative relationships between terms. By controlling the vocabulary in such a way, the process of taxonomy should eliminate ambiguous terms and explicitly show hierarchical relationships. The result is a specific and unambiguous mapping of terms.

Biological Taxonomy

Taxonomy in the biological sciences has been around for millennia. As early as Aristotle (384–322 B.C.E.) and his successor Theophrastus, one can see a clearly ordered vocabulary attempting to differentiate between species of plants and animals. Based on observations, they described in detail similarities among like specimens. Their approach was to inductively reason that which was an organism's *nature*, or that which is necessary for that organism to be identified as that organism, through observing many specimens of that organism. Aristotle organized his system from simple to complex forms of life. Aristotle and other early empirical scientists did not have the scientific tools, such as the microscope, to observe minute forms of species. Still, Aristotle's system laid the basis for taxonomic work until the 19th century, when scientists began to favor an evolutionary approach to classification as opposed to a natural approach.

As biological and medical sciences grew in the 1500s and universities began to revive a focus on studying nature, Aristotle's system was built upon or modified. Works on plants and animals showed little consistency of nomenclature, and often many different names were ascribed to a particular organism so as to describe it; names were not standardized. Generally working from Aristotle's system, Linnaeus developed a system with two kingdoms (one for plants and one for animals) and several ranks. Linnaeus, whose works were extremely detailed in description of the nature of organisms, also began to standardize names for plants and animals by developing a two-name convention system. The nomenclature, then, was separated from the taxonomy, so as to achieve a standard system. This system still remains in use today.

Since Linnaeus, the science of taxonomy has continued to develop. Two major developments have occurred. The first was an extension of the system by applying different ranks and nomenclature. The second reformed the way taxonomy was done on a more basic level, that of phylogenetic development rather than natural affinity.

The systems that developed out of the work of Linnaeus continued to shift on the basis of what was observed in the world. At present, for example,

there are two major competing systems, the domain system and the kingdom system. The distinction centers on whether all organisms can be categorized into three domains (Archaea, Eubacteria, and Eukaryota) developed by the microbiologist and physicist Carl Woese in the 1990s, or simply the five kingdoms (Monera, Protista, Fungi, Plantae, and Animalia) proposed by Robert H. Whittaker in 1969. The rank of domain was added chiefly to accommodate the discovery of the unique features of Archaea, which were difficult to categorize into one of the five kingdoms and were distinct enough to perhaps require a new hierarchical category. Eight major ranks have also been developed and generally accepted (domain, kingdom, phylum, class, order, family, genus, and species).

The second major development since Linnaeus speaks to the onset of a new worldview: that of evolutionary biology. Taxonomy is not simply a classification of *like* and *unlike* organisms but also presents a theory of the relationships between them. Whereas previous taxonomists had assumed that all living things had a nature, and thus should be categorized based on inductive observation, with the understanding of evolution, the phylogenetic development of organisms was emphasized above the direct observation. Thus, reference to the fossil record became a major strategy for the classification of species. This ultimately shifted the assignment and interpretation of classifications. Hierarchical relationships could be understood not only as a taxonomy of natural features but also to some extent a description of development.

Process of Taxonomic Classification

The approach to taxonomic studies is not necessarily standardized, but it requires first and foremost the collection of suitable specimens. Once one can collect a quality specimen of a particular type of plant or animal and then begin to compare it among the variation that already exists, detailed notes and descriptions are made to compare and contrast with what is known. If one suspects the specimen is a new or undocumented species, then a suitable name, derived from the internationally recognized codes, would be applied. Once a specimen is documented and named, it must be situated within the existing categorization structure.

The existing classifications may be revised if the specimen lends new evidence that seems to contradict the current classifications. This process, though not a standard set of procedures, is the general process of doing taxonomic work.

International and Universal Coding Systems

In order to create a taxonomy that is broadly applicable enough to be used globally, there must be some way to regulate the language. International universal codes have been published for such purposes. These codes define a standard nomenclature and classification system, but the content, of course, is left to taxonomists who apply the system to real specimens.

The International Association for Plant Taxonomy, for example, publishes their *International Code of Nomenclature for Algae, Fungi, and Plants*. This document dictates specific word endings for ranks, such as *-phyta* for any plant phylum. As such, this document serves as a control to allow for the naming, classification, and cross-referencing all plant, algae, and fungi organisms. These types of universal codes allow for the standardization of a taxonomy that incorporates methodological rigor and peer review of data. They also ensure the standardization of naming. As consistent with the long conventional history rooted in Linnaeus, the naming of any species will be unique, including a binominal name made up of the genus and species name. Sometimes, it is important to distinguish the subspecies, as if one was referring to regional differences in a species of tree. In these instances, a trinomial name consisting of the genus, species, and subspecies would be used.

Examples of Taxonomic Names

The standard classification of species, as noted above, usually contains eight major ranks. This allows for hierarchical specificity. However, the name of a particular species usually contains only the genus and the species name. Below is the taxonomy of the modern human, the cougar, the mountain ash from the example above, and the white ash to show the hierarchical distinction.

Table 1 Examples of Official Nomenclature

Rank	Modern human	Cougar	Mountain ash	White ash
Kingdom	Animalia	Animalia	Plantae	Plantae
Phylum	Chordata	Chordata	Magnoliophyta	Magnoliophyta
Class	Mammalia	Mammalia	Magnoliopsida	Magnoliopsida
Order	Primates	Carnivora	Rosales	Scrophulariales
Family	Hominidae	Felidae	Rosaceae	Oleaceae
Genus	<i>Homo</i>	<i>Puma</i>	<i>Sorbus</i>	<i>Fraxinus</i>
Species	<i>H. sapiens</i>	<i>P. concolor</i>	<i>S. aucuparia</i>	<i>F. americana</i>
Subspecies	<i>H. s. sapiens</i>			

Each scientific name is binomial, though sometimes the trinomial name may be used. For example, modern humans are *Homo sapiens*, but in order to differentiate them from their immediate predecessors, the trinomial name *Homo sapiens sapiens* is used. A predecessor, though the same species as modern man, has been distinguished through fossils in Eastern Africa and named *Homo sapiens idaltu*. The distinction is still up for debate as fossil records are updated.

The species and genus name make up the official nomenclature. Some extend back to Linnaeus's classifications.

As can be seen from the examples in Table 1, the scientific specificity of each species allows botanists and zoologists to cross-reference species with their hierarchical relatives. Any organism would be identified through this taxonomy.

Other Applications of Taxonomy

Among the various taxonomies that are used on small and large scales, some have become quite widely accepted. One that has been very influential in education is the taxonomy of learning domains, developed by the educational psychologist Benjamin Bloom. Bloom's taxonomy of the cognitive domain was part of a larger work that included three categories that could be useful in training or education: *cognitive*, *affective*, and *psychomotor*. Within these categories, different abilities or skills can be identified. Within the cognitive domain, for example, six categories may be identified. These have been extended and revised to include six verbal categories: *creating*,

evaluating, *analyzing*, *applying*, *understanding*, and *remembering*. This taxonomy is broadly applied in education, as many teachers are instructed to write lesson plans that focus on more than a single domain and to ask questions to students that focus on different cognitive domains. Thus, the taxonomy has provided a sort of template to think about cognition in the classroom, in assessments, and in education more generally.

Abraham Maslow, a noted psychologist, developed another widely used taxonomy. His hierarchical schema of what motivates humans can be understood as a taxonomy. It categorizes the needs of humans from the most fundamental, such as breathing and eating, to more complex needs for self-actualization. This theory is highly influential in psychology, although it has drawn some criticism on the grounds of its empirical underpinnings. Nonetheless, this hierarchy, as with many different developmental theories, attempts to taxonomize human development by standardizing the way we categorize its aspects.

Joseph C. Rumenapp

See also Life Sciences, Contemporary; Life Sciences, Historical

Further Readings

Anderson, L. W., Krathwohl, D. R., Airasian, P. W., Cruikshank, K. A., Mayer, R. E., Pintrich, P. R., Rath, J., & Wittrock, M. C. (2000). *A taxonomy for learning, teaching, and assessing: A revision of Bloom's taxonomy of educational objectives*. Pearson.

- Atran, S. (1993). *Cognitive foundations of natural history: Towards an anthropology of science*. Cambridge University Press.
- Bloom, B. S. (1956). *Taxonomy of educational objectives, Handbook I: The cognitive domain*. David McKay.
- Durkheim, É. (2008). *The elementary forms of religious life* (M. Cladis, Ed. & C. Cosman, Trans.). Oxford University Press.
- Lamberts, K., & Shanks, D. R. (1997). *Knowledge, concepts, and categories*. MIT Press.
- Linné, C. von, & Salvius, L. (1758). *Caroli Linnaei Systema naturae per regna tria naturae: secundum classes, ordines, genera, species, cum characteribus, differentiis, synonymis, locis*. Holmiae: Impensis Direct. Laurentii Salvii. <http://www.biodiversitylibrary.org/item/10277>
- Maslow, A. H. (1943). A theory of human motivation. *Psychological Review*, 50(4), 370–396.
- McNeill, J. (2012). *International code of nomenclature for algae, fungi and plants (Melbourne code)*. Koeltz Scientific Books.
- Whittaker, R. H. (1969). New concepts of kingdoms of organisms. *Science*, 163, 150–160.
- Woese, C. R., Kandler, O., & Wheelis, M. L. (1990). Towards a natural system of organisms: Proposal for the domains Archaea, Bacteria, and Eucarya. *Proceedings of the National Academy of Sciences of the United States of America*, 87(12), 4576–4579.

THEORY CHANGE

Science strives to develop theories that explain its subject matter of *reality*. When a theory passes severe tests, when its scope can be expanded, then scientific knowledge-seeking is deemed to be successful. But if a theory changes or is even replaced by a new theory, if it emerges that a particular theory has possibly been wrong, then mistrust may attach to the scientific enterprise. Theory change has thus become a double-edged sword at the heart of the self-reflexivity of science that has engendered an intense and ongoing debate in the philosophy of science. This entry traces the origins and emergence of the shift from an understanding of theory as logical analysis to one that takes into account the embeddedness of knowledge in history. Among the key topics covered are 20th-century developments such as non-Euclidean geometry, relativity, quantum mechanics, logical empiricism,

and critical rationalism, and the conception of theory change as radical shift as opposed to an incremental rational correction guided by logical rules. The contribution of societal, extrascientific factors, the concept of paradigms, and the contribution of experimental and cognitive practices are described and explained.

In what follows, we will examine how theory change, on the one hand, emphasizes the *historicity* of science. A historical product can change, and such change implies the transformation of a given *worldview* into a new one. A coherent system of *seeing* the subject matter of a scientific discipline is replaced by a different system. Theory change rejects the commonsense view of knowledge-seeking as pure accumulation of facts. Not only single elements of knowledge but the whole picture can change. Instead of comprehending science as a cumulative enterprise, theory change therefore emphasizes the active side of science: theory as something that is constructed. Consequently, scientists cannot only be concerned about the collection of facts and about how their discoveries should be organized under a given system of thought; scientists must also explore why and how new theories arise that shed new light on their subject matter. In this sense, theory change has raised previously unasked but nonetheless important questions. On the other hand, at the same time, the very same idea has also brought new doubts into science, the seed of relativism. That is, theory change calls into question the *rationality* of science by asking to what extent theories, which are only temporarily valid because they can possibly change, can still be considered objective.

It is interesting to see that theory change became a theme of discussion in the philosophy of science only at a relatively late date. It did not receive any attention at all in the first half of the 20th century. In the same package as related concepts such as scientific change and conceptual change, theory change was first abruptly and intensively discussed only in the 1960s but subsequently moved up the agenda of philosophy of science to become one of its most central and important themes in the second half of the 20th century.

For this reason, it is important to see, first, why and how in the middle of the 20th century an established understanding of theory was suddenly and successfully replaced by a new one.

The notion of scientific theory as a finished product that can be rationally reconstructed by means of *logic* was replaced by an understanding of theory as a *historical product* of human activity. The result was that the focus moved from theory as a static given to the constant possibility of theory change. Second, the terms and problems of the new understanding of theory change as a *radical shift*—a concept introduced first and foremost by Thomas S. Kuhn—have to be analyzed. Third, it will further be interesting to learn which new questions this has subsequently raised and which new answers have been given, by examining recent approaches that focus on the experimental and cognitive *practices* in the *process* of theory change.

From Logical Analysis to Historical Change

The conflictual understanding of theory as either a logical form or as historically changing first featured in the debate between the German early modern philosophers Immanuel Kant and Georg W. F. Hegel. When Hegel criticized Kant's epistemology in the early 19th century, the same problem was already at stake. Kant had powerfully articulated the constructionist idea of modern thinking that knowledge is a human product. While attempting to provide a philosophical foundation for Newtonian physics, Kant reasoned that the scientific understanding of the world cannot be *anchored* in empirical sensations or perceptions. He rejected the empiricist belief that knowledge can be acquired by simply aggregating experiences. In contrast, he postulated that knowledge is shaped by cognitive structures. Kant claimed that forms of intuition (space and time) and categories (such as substance or causality) are transcendental forms of a logical subject that must be assumed for understanding the acquisition of knowledge. Accordingly, knowledge is possible only within a distinctive frame of forms and categories. In short, Kant defends the priority of concepts over facts. Theory gives form and shape to reality, which for Kant is otherwise chaotic. Because the knowledge about the world is mediated by logical categories, Kant's approach did not yet raise the problem of relativism. For Kant, the objectivity of knowledge can still be

justified insofar as theory, the frame of reference, is anchored in transcendental structures of the knowing mind that are permanent and universal. Hegel, and after him other thinkers of the 19th century, rejected precisely the idea of any form of eternal truths that are immune to historical change. For Hegel, science cannot be based on transhistorical forms of knowledge. It cannot be based on a logical subject. On the one hand, Hegel shares with Kant the idea that knowledge is constructed via a categorical framework. On the other, and in contrast to Kant, Hegel historicizes this framework. He *de-transcendentalizes* the knowing mind—Kant's logical forms and categories—by locating reason in historical time and social space. In this sense, Hegel introduced history to epistemological analysis.

In sum, Kant and Hegel both stress the importance of theory in the construction of knowledge. However, whereas Kant still assumes a static world by claiming the dependence of thought on intuition, Hegel's phenomenology of different forms of knowledge emphasizes the embeddedness of knowledge in history. At the turn of the 19th century, Hegel introduced the idea of change to the understanding of theory: Even theories are historical entities; they are revisable. Theories change in history.

By the early 20th century, Hegel's influence had waned and Kant's atemporal philosophy had been strongly undermined, ironically by a factual, that is, historical theory change in science. The invention of non-Euclidean geometry in the 19th century and the uncertainty principle in quantum mechanics in the early 20th century required a revision of Kant's foundationalist principle of the a priori limits of scientific knowledge. In Kant's time, the worldview of a Newtonian clockwork mechanism underlay the scientific understanding of the universe. In Newtonian physics, the Euclidean metric appeared unshakeable. It was an axiom that provided a priori knowledge about the relations between space points, that is, about the structure of space. Kant accordingly supposed Euclidean geometry to be the repository of a priori truths that scientists impose on what they perceive. Since the emergence of non-Euclidean geometries in the 19th century and Albert Einstein's theory of relativity in the early 20th century, declaring that

physical truth is relative to geometric form, the actual geometry of the scientific worldview is no longer Euclidean. Accordingly, the dominant philosophy of science of the first half of the 20th century—logical empiricism—which framed the debate for several decades and was strongly rooted in the relativistic and quantum revolution in physics—could also no longer ground science in a priori truths about the natural world. The logical empiricists of the Vienna and Berlin Circles of the 1920s and early 1930s therefore shifted the primary focus from the quest for epistemological truths grounded in metaphysics to grounding epistemology in scientific methodology itself. Two strategies became central: an emphasis on empirical evidence and the application of logical analysis. First, only what is immediately given by the senses can be known for certain. No knowledge is immune from revision in the light of experience. Correspondingly, the limits for the content of science are set. Only *empirical statements* can be relevant. Unobservables are either deleted from the scientific agenda or they must be associated with terms for observables. Consequently, the manifesto of the Vienna Circle group proclaimed that any term in science must be ascertainable by stepwise reduction to basic terms of sensory experience. Besides propositions that express mathematical or logical truths, there are either observation statements or propositions that are meaningless because there is no observational evidence for them. Furthermore, observational facts, such as the existence of things of all kinds or experimentally observed correlations between variables, are differentiated from their theoretical interpretations, for instance, the concept of a gravitational force or the kinetic theory of gases. Facts and theories are sharply distinguished. If theory is necessary at all, it is secondary to the facts, given that the primary aim of the logical empiricist methodology is to base science on a nontheoretical, observational language. Second, in order to further keep all forms of metaphysics out of science, the only legitimate method to be applied to the hierarchical order of terms is *logical analysis*, that is, movement from elementary or protocol statements on the lowest level of actual experience to further statements via a controlled procedure. For Rudolf Carnap and other Vienna Circle thinkers, *inductive logic* guides the path from the most basic empirical statement of

immediate sense-data (*this tomato is red*) to other statements, such as empirical generalizations (*tomatoes are red*). Any hypothesis such as *tomatoes are red* can thus be true or false, but it cannot be meaningless; it will be either confirmed or invalidated by an accumulation of basic protocol sentences. Accordingly, theoretical terms such as *electron* do not refer to tiny particles but to logical constructions based on experiences. Both strategies—the reduction of theoretical interpretations to records of experience and the corresponding emphasis on logical analysis—had drastic consequences for the understanding of the possibility of theory change. Attracted by the rigor and clarity of mathematical logic, the attention of the logical empiricists was drawn away from change and directed at atemporal components of the scientific toolbox such as laws and confirmation. If theory change is a useful concept at all, it has to be understood as an accumulative process because the basics of science, the empirical statements, do not change. Generally speaking, science is the accumulation of empirically certified truths. Theories that replace each other are therefore also linked to each other. Any former theory can be reduced to a succeeding theory; that is, the former can be considered a special case of the latter. Science is continuously growing, step by step. There is progress in science, with each successive stage being more stable than its predecessors.

This view, based on the strict separation of the logical and the historical, was further supported by a widely accepted and very influential distinction made by Hans Reichenbach, who was the founder of the Berlin group of logical empiricists. Reichenbach distinguished between two realms of the scientific enterprise, placing them strongly in opposition to one another: the context of discovery and the context of justification. The latter refers to the logical analysis of scientific knowledge, while the former regards questions concerning the origin or generation of scientific ideas. Whereas all issues around the question as to how scientists arrive at their knowledge claims are therefore considered to be tasks for psychology, sociology, and history, the proper domain of the philosophy of science, according to Reichenbach's distinction, is the epistemic status of a claim, that is, the evaluation of scientific knowledge. Philosophy of science thus does not consider real

processes of knowledge acquisition. Instead, the task of epistemology is to evaluate the correctness of knowledge claims by arranging them in a consistent system, that is, the legitimacy of the reasons, arguments, rules and standards behind the empirical testing, and thus the confirmation or disconfirmation of knowledge claims will be considered. By contrasting the context of discovery and the context of justification of knowledge, it is obvious that, for example, a discussion about how views on the structure of space are acquired cannot help to ascertain whether Euclidean or non-Euclidean geometry of space is more justifiable. Following the neo-Kantian distinction between genesis and validity, logical empiricism separated logic from history. Instead of being interested in the diachronic dimension of theory change, the dominant approach in the philosophy of science in the first part of the 20th century focused on the synchronous investigation of the finished products of scientific research, defined as the mathematical and logical structure of (mainly) physical knowledge.

It is interesting to see to what extent the work of another influential philosopher of science, Karl R. Popper, who founded critical rationalism, the second major professional philosophy of science in the 20th century after the logical empiricist movement, supported or undermined this established view. Although Popper studied and received his PhD in the late 1920s in Vienna, he was never invited to become a member of the Viennese group of logical empiricists. On the one hand, both the logical empiricist and the Popperian approaches claimed that discovery has no place in the logic of science. That is, according to Popper, it is simply irrelevant to the logical analysis of scientific knowledge on how new ideas are received or created. On the other hand, although sharing with the empiricists the emphasis on experience, the view that theories can only be called scientific if their validity can be asserted in the light of observation or experiment, Popper also strongly deviated from the logical empiricist focus on verification in scientific methodology. For Popper, science does not begin with observations, perceptions, or *brute facts*. For this thinker, science begins with problems. Popper sees science as a problem-solving enterprise. Experience cannot determine theory and thus science cannot infer

theory from observational statements. It cannot inductively arrive at generalizable universal propositions because science does not proceed inductively but deductively. First, a new theory will be logically analyzed; its internal consistency will be tested. Then the theory in question will be compared to existing theories. If it does not promise greater empirical depth than its forerunner, it will not be adopted. If it promises to solve previously unsolvable problems, it will be adopted. Finally, the crucial step is its empirical testing. Theories contain logical and empirical elements. More specifically, theories must be operationalized so as to suggest hypotheses in circumscribed research areas. Hypotheses guide observations, where observations refer to perceptions. In other words, science does not begin with perceptions. Perceptions are actively planned and prepared through observations, which themselves are also selective; they are always preceded by something theoretical: questions, problems, or hypotheses. Popper's *searchlight* model of knowledge acquisition therefore explicitly opposes the *bucket theory* of knowledge, as he termed the empiricist approach. Science does not proceed by collecting facts; science proceeds by testing the empirical application of conclusions derived from theories. If a theory stands the test, it can be maintained. If it does not stand up to the attempt to refute it, it has to be abandoned. For Popper, therefore, falsifiability, not inductive verification, demarcates science from nonscience. The idea that all knowledge is provisional makes science a fallible enterprise. Scientific theories cannot be proved, they can either be provisionally confirmed or conclusively refuted. Although the growth of knowledge in science can accordingly no longer be understood as a cumulative process but must be seen as an endless procedure of error elimination, for Popper's theory change in science nevertheless involves theoretical progress because a newer theory must simply have greater empirical content than its predecessor. For example, the gravitational theory of Isaac Newton has had greater generality than Johannes Kepler's and Galileo Galilei's laws of motion. According to Popper, the former cannot be considered as having been inductively inferred from the latter theories. On the other hand, the explanatory scope of Newton's theory not only included both Kepler's kinematics of celestial

motion and Galileo's kinematics of terrestrial motion; Newton's dynamics also provided a numerical measure that was close to, but different from, the theories of Kepler and Galileo. The same then applies to the replacement of Newton's theory by Einstein's theory of relativity in the early 20th century: Einstein's theory cannot be inferred inductively from Newton's; the two can be compared only by deductively testing them empirically and by demonstrating that one explains more than the other.

For the logical empiricists, theory change was not a matter of concern. They saw the body of knowledge as continually growing and theories as nothing more than snapshots in an ongoing learning process. For Popper, in contrast, theory change is no longer a cumulative process. It is a trial-and-error process with the objective of error elimination. But despite the strong theoretical conflicts between the two philosophies of science that were dominant in the 20th century until the 1960s, they also shared certain characteristics. On the one hand, they differ regarding the primacy of perception versus selective observation and on the role of induction versus the deductive testing of theories' scientific methodology. On the other hand, they share the emphasis on logical analysis. In both approaches, it is the rigor and clarity of a formal method that guarantees continuity and intersubjective validity. For both, science is a rational enterprise that advances by increasing approximation to truth. Although Popperians presuppose the fallibility of theories, for them falsification qualifies as the logical technique guaranteeing that any current theory can be provisionally maintained as the theory of greater empirical depth than its predecessors, that is, as the best available theory—until it is falsified. The critical rationalist method thus invites theory change as a change toward a *better* theory, which is distinguished by its capacity to solve more problems in the domain of the scientist than competing theories do.

For both methodologies, theory change is still under rational control. Logical empiricism and critical rationalism both share a general *logical* approach insofar as for the latter the logic of justification lies in falsification. They do not yet regard theories as historical entities that can be replaced by completely different theories for contingent historical reasons. By contrast, the

consideration of theories as historically given invites a new understanding of theory change. The new view takes into account the possibility of the subject matter of the scientific investigation also shifting in the process whereby one theory is replaced by a newer one. If this is so, if even the meaning of terms can accordingly change from one theory to its successor, then theory change implies a radical discontinuity in the development of knowledge.

Theory Change as Radical Shift

If theory change implies a radical shift as opposed to an incremental rational correction guided by logical rules, then a succeeding theory will no longer be understood as an improved version of the same theory but as a radically new theory. Examples are Newton's notion of gravity, Maxwell's dynamic theory of the electromagnetic field, Mendel's theory of genetics, and Einstein's theory of special relativity, all of which marked a deep transformation in the scientific understanding of nature.

To introduce the possibility of theory change as historical project in this radical new sense, new philosophical sources became important in the mid-20th century. Ludwig Wittgenstein's theory of meaning provided the main philosophical toolbox in this context. For Wittgenstein, meaning is not a separate entity. It is neither a mental image nor does it represent some object in the world. The meaning of a word depends on the uses made of it in various activities of human life when the users of a language are exposed to actual situations in which they must apply the word appropriately according to the rules and grammar of language. Understanding meaning in this way, not as a given entity but as a product of an actual *game*, that is, as a relational result, will be helpful for further understanding the arguments made against the philosophy of science mainstream in the mid-20th century. The attack that rang the knell for empiricism and the associated gradual interpretation of theory change was finally carried out by Thomas S. Kuhn's epoch-making book *The Structure of Scientific Revolutions*. The aforementioned arguments by other authors, however, prepared the ground for Kuhn's intervention. Basically, there were two complementary arguments.

First, in the late 1950s, the thesis of *theory-ladenness of observation* simply denied the existence of theory-neutral data. On the one hand, this thesis argues, theoretical presuppositions held by the investigator affect even perception. When a Ptolemaic believer in the Earth-centered universe and a Copernican believer in the sun-centered universe are watching a sunset, they see different things. The Ptolemaean sees the Earth to be stationary and the sun to be falling, whereas the Copernican sees a moving horizon and a fixed sun. Accordingly, people, not eyes, are seeing, as Norwood Russell Hanson, one of the authors who elaborated the thesis, concisely summarized it; and so for Hanson, even seeing is a theory-laden undertaking. Beyond perception conditioned by background assumptions, on the other hand, there is another version of theory-ladenness, namely, semantic theory-ladenness. According to the latter view (and as for Wittgenstein), scientific concepts have meanings that cannot be reduced to immediate experience. For example, the observational term *mass* has a different meaning in Newton's physics than in Einstein's theory of special relativity. Whereas in previous physical theories, mass was viewed as constant, in the theory of special relativity, mass and energy cannot be viewed or measured as distinct entities. As long as theories partially determine the meaning of observational terms, two investigators attached to two different theories will mean different things even though they might be applying the same observational terms. If theory through which facts are interpreted matters as thoroughly as postulated here, then the formerly upheld empiricist theory–observation dichotomy must be abandoned.

Second, the complementary argument that has further undermined the distinction between logical analysis and experience is the thesis of the *underdetermination of theory by data*. While according to the theory-ladenness thesis, science does not have access to pure *facts* not affected by theories, the theory–observation relation can also be considered from the theory point of view. Considering again the aforementioned example, it could be asked whether the Ptolemaic or the Copernican *theory* should be preferred, although a widely held view in Copernicus's time was that the two rivals did not actually differ in their predictions of the available astronomical data

(although they differed strikingly with respect to the unobservable). In sum, a particular theory is underdetermined by observational evidence. The difficulty has been fueled even further by a complementary argument offered in the 1950s by Willard V. O. Quine, also known as the *Duhem–Quine thesis*. Quine made the holistic assumption that theories can only be tested against evidence as a whole. Experience cannot confirm or falsify individual beliefs but only the coherent set of beliefs that had been fit together to form a theoretical whole. That is, on the one hand, hypotheses cannot be tested individually; on the other hand, any aspect of a theory can be retained by adjusting auxiliary hypotheses elsewhere.

Both of these arguments supported the outstanding success of Kuhn's landmark book—and vice versa. Kuhn's book, first published in 1962, changed the course of philosophy of science in the 20th century, including its understanding of scientific change and, more specifically, of theory change. Before Kuhn, science was understood to be a linear, cumulative unveiling of reality; theory change was not a point of discussion. According to the Google Books Ngram Viewer, *change*—whether scientific change or theory change—was not an issue before the 1960s. Before Kuhn intervened, the further development of a theory was understood as a cumulative growth of knowledge by either accumulation of observational evidence or error elimination, which, according to Popper, supports theoretical progress (since every succeeding theory contains the positive results of the earlier theories). Kuhn changed everything. Since Kuhn, theory change no longer means growth. Moreover, it no longer implies a rational process. Since the 1960s, theory change may involve change in the meaning of the terms used in two succeeding theories; it may therefore represent a radical break. Kuhn's immensely influential book introduced three key concepts: *paradigm*, *scientific revolution*, and *incommensurability*.

First, it is neither observational statements nor isolated hypotheses that matter first in science, rather it is *paradigms*. A paradigm, later also called a *disciplinary matrix* by Kuhn, is more than just a theory. In accordance with the holistic viewpoint outlined above, a paradigm includes a set of background assumptions. It is a deeply anchored general framework that serves several functions at

once. It defines a problem field, it offers a source of methods, and it contains the standards of solution, which are endorsed by all scholars working in the relevant scientific domain. It is most definitely not easy to change. Where Popper saw refutations when theories faced difficulties, Kuhn sees anomalies against which paradigms are immune. That is, where the Popperian scientist has to falsify her hypothesis, the Kuhnian scholar would never give up the paradigm on which all research in her field is based. Consequently, theory change along this line would no longer guarantee scientific progress. The reason is that, according to Kuhn, there are further sources beyond empirical observations and tests that contribute to establishing and keeping a paradigm at work.

Second, once a paradigm does actually change, the scientists concerned will indeed be working in a different world. This happens in *scientific revolutions*. In long periods of *normal science*, the daily work of the scientist could also be described in terms of the above-mentioned philosophies of science. Scientists put the results of their studies together to form a gradually growing body of knowledge. All of them work under the same paradigm. Normal science has to be considered as a conservative period in the history of science. Scientists know and share the rules of the *game*. They simply but ably apply these rules to new phenomena. Contrary to the popular view, scientific research in normal periods of science does not really *discover* anything new, at least not a new view on reality. In actual fact, all scientists share the same view of reality. This changes in revolutionary periods. In *scientific revolutions* such as the 16th-century Copernican revolution in astronomy, the 19th-century Darwinian revolution in biology, and the 20th-century Einsteinian revolution in physics, confidence in the existing paradigm breaks down. A sense of crisis develops in the scientific community. Anomalies that cannot be accommodated in given theories are usually ignored. In revolutionary periods, however, once a growing number of anomalies have accumulated, more and more scientists become open to accepting a new paradigm. In contrast to the gradual growth of knowledge during the period of normal science, a scientific revolution involves a radical shift. According to Kuhn, Galileo, for example, lived in a different world from Aristotle.

Third, the radical-shift thesis is implied in Kuhn's thesis of *incommensurability*. If experience depends on theory, the problem arises that two scientists who hold different theories will mean different things even if they use the same terms to describe their observations. In Ptolemaic astronomy, *planet* referred to the Sun and the Moon but not to Earth. In the Copernican worldview, these meanings have changed. The sun has moved to the center of the solar system, and the Earth and all other planets are thought to revolve around it. If two competing theories are incommensurable in this sense, they are even incomparable, as both standards of comparison, including meanings of concepts, are dependent on the two different paradigms to which they belong. In summary, if theory change is understood as a Kuhnian radical shift, it produces incommensurable conceptual languages and different ontological worlds. Theory change can then no longer be considered a rational enterprise.

Theory Change as Practice

Whereas earlier philosophy of science did not study the *context of discovery*, Kuhn's holistic-historical understanding of theory made philosophers aware of the role of extrascientific factors in theory change. By no longer focusing on rationality in logical extensions and the relation between terms and empirical phenomena, Kuhn's approach invited the study of factors such as the interests of involved scientists, governments, funding agencies, or the influence of technology or gender. Since the 1970s, this kind of research has been carried out extensively by scholars of the sociology of scientific knowledge (SSK). For SSK, science is a social activity; it is practice. Theory change then is no longer regarded as a rational process but as a social enterprise.

While both SSK and Kuhn's view can be understood as the antithesis to the previously dominant logical analysis of scientific knowledge, Kuhn also opened doors to other approaches that take a third path. The study of both *experimental* and *cognitive* practices is of major importance in this regard. Both share an interest in actual scientific activity. Instead of simply comparing two completed conceptual structures, they focus on the *process* of theory change.

First, Ian Hacking's *Representing and Intervening* claims that experiments have a life of their own that is independent of the representation of their results. When scientists use microscopes, they are not applying a theory; they are not just *looking*, they are *doing*, they are manipulating things. Even if different microscopes are used, often the same patterns are found. In this respect, theories are preceded by stable practices. Hacking's view thus makes an important point against the *incommensurability* thesis. Whereas a Kuhnian researcher focuses on that which varies (and then claims that this involves two incompatible worldviews), Hacking's practice-oriented account of theory change considers that which does not vary. In this latter perspective, there is something that is preserved even across theory change that comes to light in experiments. Consequently, Hacking holds a *realist* view of the process of science according to which theory change is not to be regarded as irrational, as it would appear in the light of Kuhn's relativism. While Kuhn denied the existence of a third, commonly shared framework that would be needed to enable the comparison of two succeeding theories, the successful manipulation of things in experimental practices shows that there must be something in reality that corresponds to the theories presupposed in the analysis.

Second, the study of the *cognitive* processes offers a further approach that does not limit itself to the comparison of end products in the process of theory change. Instead, by studying epistemic practices that scientists apply in problem-solving processes, such as analogies, visual models, and thought-experiments, the cognitive approach to science shows how piecemeal transformations of theoretical frames can result in revolutionary changes. For example, Lindley Darden has suggested focusing on the reasoning strategies that could have produced the theoretical modifications of Mendelian genetics between 1900 and 1926. One of the strategies for changing theories that she found is the strategy of *delineate and alter*. It involves the analysis of a component into several parts and then a change to only some of those parts. Another approach that reestablishes reason at the core of theory change has been offered by Paul Thagard, who makes a plea for a *cognitive*

science of science. Thagard proposes to analyze the conceptual system that makes a theory by focusing particularly on kind-hierarchies and part-whole relations in the organization of concepts and then considering which computational mechanism could have produced the components of the conceptual system. His analyses have shown that explanatory coherence is the core mechanism of theory adoption in conceptual change.

While the cognitive approach to theory change moves the focus to a fine-grained analysis of conceptual developments, it still assumes that scientific concepts are embedded in theories that are revisable over time. It thus shares the basic assumption of other forms of historical analysis against the synchronic investigation of the end products of scientific change that there are other rules at work (and to be explored) in the process of theory change, beyond the reference to observational evidence (and its rational reconstruction).

Frank Welz

See also Scientific Revolutions; Social Construction of Scientific Knowledge

Further Readings

- Darden, L. (1991). *Theory change in science: Strategies from Mendelian genetics*. Oxford University Press.
- Hacking, I. (1983). *Representing and intervening: Introductory topics in the philosophy of natural science*. Cambridge University Press.
- Hanson, N. R. (1958). *Patterns of discovery: An inquiry into the conceptual foundations of science*. Cambridge University Press.
- Kitcher, P. (1993). *The advancement of science: Science without legend, objectivity without illusions*. Oxford University Press.
- Kuhn, T. S. (1962). *The structure of scientific revolutions* (2nd ed.). University of Chicago Press (with a new "Postscript".1970).
- Pickering, A. (Ed.). 1992. *Science as practice and culture*. University of Chicago Press.
- Popper, K. R. (1959). *The logic of scientific discovery*. Hutchinson & Co.
- Thagard, P. (2012). *The cognitive science of science. Explanation, discovery, and conceptual change*. MIT Press.

THEORY CONSTRUCTION

Theory construction is a process through which propositions are developed, tested, scrutinized, and revised to explain observations for which no reliable previous explanation exists or for which existing explanations are flawed or incomplete. Several steps are involved in theory construction across all fields of study that enable new propositions to arise while ensuring that those propositions are accurate across a wide range of contexts. In the following sections, each step will be described and discussed in turn.

Theory Origination

Theory construction begins with an observation for which the theorist has no verifiable explanation but does have concepts or constructs that enable recognition of the phenomenon and the generation of a question or prediction about what is observed. Concepts—for example, *shirt* and *jacket*—are concrete ways of labeling phenomena, whereas constructs, such as *style*, are more abstract applications of concepts. *Thought trials*, as Karl Wieck (1989) describes them, are means through which theorists use concepts and constructs to construct questions about observations and consider possible solutions.

Concepts and constructs, in turn, are informed by paradigms that shape theorists' ontological perspectives. Paradigms are general ways of viewing reality that frame how theorists make observations, devise questions, and generate meaning from findings. For example, to some degree, East Asian health sciences operate within a somewhat different paradigm than European or American health sciences, as a result of historical forces having shaped the ways that illness has been conceived and medicine practiced in different regions.

Paradigms fuel theory construction by suggesting what is important to think about, to know, and to investigate. Also, paradigms guide theory construction by suggesting what relationships should exist and what the observable effects of those relationships are likely to be. At the same

time, paradigms can also limit theory construction by conditioning what is perceived, what is thought worthy of exploration, and what sorts of questions are asked.

In the process of theory development, originating paradigms might be found to be faulty if expected relationships and effects found by researchers are absent or different from what is expected within the paradigm. Entire paradigms can be rejected and replaced because they can no longer generate theories that reliably explain observations. Thus, what Thomas Kuhn described as *scientific revolutions* can be the result of theory generation initially inspired by the very paradigms that are subsequently discarded. A classic example is the Copernican Revolution of the 16th and 17th centuries, which replaced the Ptolemaic paradigm of the 2nd century C.E. with a new view of the universe wherein the Earth revolves around the sun rather than vice versa. This shift in paradigmatic thinking made old theories obsolete but opened the way for new ways to explain the universe.

Forms of Reasoning

Theory generation can be either a *deductive* (sometimes called *top down*) reasoning process or generated *inductively* (*bottom up*). When constructing theories deductively, theorists begin with a principle that establishes a horizon of possibilities within the respective paradigm. That principle is then applied to specific observations to refine a singular explanation that will hold true across a wide range of instances. Findings derived from deductive approaches either support the initial proposal or prove the proposal to be false. Theorists must keep an open mind, therefore, to the possibility that new observations might contradict established principles and thus challenge the established paradigm.

Theories generated inductively begin with specific observations in which theorists look for patterns among phenomena of interest. Established theories can be used to explain the patterns, or new theories could be devised where no previous explanations suffice.

A third possibility, less discussed in literature on theory construction, is *abduction*. In this case, unexpected results lead to insights that might not have been conceivable via inductive or deductive approaches. Abductive reasoning relies upon inferences made about causes of unexpected outcomes.

Division of reasoning into three categories for heuristic purposes is, however, somewhat misleading because induction, deduction, or abduction are not entirely separable strategies in theory construction. *Grounded theory*, as proposed by Glaser and Strauss, for example, combines induction and deduction via *the constant comparative method*. Using this method, the researcher moves from observations of specific instances to tentative explanations for those patterns and then back to specific instances to test the tentative explanations. Similarly, abduction requires at least some degree of reasoning based upon known principles to explain unexpected findings which then leads to further testable hypotheses.

Furthermore, as Weick points out, the step-by-step approach of theory building described thus far is an idealized, rational, linear model. In reality, the process is often interrupted and redirected by intuition, emotional response, and unplanned circumstances, inspiring insight, which in turn accelerates, slows, or redirects the theorist's progress.

Analogies are often important elements in the theory construction process, as Brian Haig has observed. Analogies permit inferential leaps of logic from the known to the unknown via comparison of phenomena that are presumably alike in fundamental ways. For example, gatekeeping theory in the field of mass communication research asserts that individuals, policies, and values within news organizations permit some stories to be published (make it through the gate), while other potential stories are not published because of gate closure.

Causation is an important, albeit somewhat controversial, element in theory construction. Theories are developed to explain if, why, and how changes in one concept—the independent variable—affect other, dependent, variables. Within controlled settings, researchers test causation by manipulating independent variables to test the accuracy of relationships expected within the operative paradigm. Beyond controlled settings, researchers can also test

relationships by examining how variables relate to one another in natural settings.

There is some debate as to whether causal relationships are an advisable form of relationship to theorize about, given the unobservable nature of causation. Instead, some scholars such as David Whetten assert that it is more prudent to describe and predict relationships that are either categorical (demonstrating an association between A and B) or sequential (ordered in time so that B is followed by A).

Hypotheses and Research Questions

Hypotheses and/or research questions are developed to predict, inquire about, or at least contextualize specific relationships among variables relevant to the overall theory. Hypotheses are specifically used to test theoretical assumptions by predicting the relationships and effects generated among selected concepts and constructs within the domain of the theory. Specific instances of more general constructs are examined to test whether expected relationships and effects are valid across multiple conditions. For example, if a theory suggests that an increase in playing violent video games leads to a decrease in hostile feelings toward others, then a resulting hypothesis might suggest that an increase in hours male undergraduates spend playing *Tour of Duty* would lead to an increase in those same students' hostility toward female undergraduates.

A research question, in contrast, is not designed to test a theory, but rather to inform inquiry about specific phenomena. Using the same example, a research question might ask about the nature of violence in video games, thus leading to a classification and quantification of types of violence present in the game. In short, research questions can function to explore relationships among variables, whereas hypotheses test for expected relationships.

Variables, in turn, can be quantifiable and thus measurable, or they can be more qualitative and conceptual. For example, the average number of hours students study per day can be measured by tallying study hours among a sample of students over a number of days and then dividing the total hours by number of days and number of students in the sample. In contrast, attitudes and beliefs

about studying are less easy to quantify but may be gauged in relative terms by the ways studying is discussed in personal narratives among students in the sample.

Once hypotheses and/or research questions are devised, a methodology is determined for collecting and analyzing data. If human subjects are involved, studies' purpose and method of data collection must be approved by institutional review board of sponsoring organizations. Data are then collected and analyzed following the planned methodological steps.

Professional Scrutiny

To build a credible theory, data observed, study findings, and the reasoning process used to interpret the data are written in a format with the intent of publication in a scholarly research journal or book. Research reports are typically organized in a patterned way with the importance of the topic stressed first, followed by a literature review summarizing what other scholars have already found regarding the topic and perhaps how other theories have been used to explain the topic. Having thus set up the intellectual context of the study, the hypotheses and/or research questions driving the study are described as is the method used to collect and eventually analyze the data. Findings are then summarized and explained in relation to the generating theory. Limitations and weaknesses of the study conclude most written reports along with new questions raised by the study.

Prior to publication, research reports are usually presented to scholars and practitioners at professional conferences for review and critique. Most often, to be eligible for presentation, scholarship must pass through a *blind* peer review process wherein two to four anonymous researchers—representing a range of expertise from graduate students through tenured faculty members—are asked to read and rate submissions. Conferences at the national level might accept as few as one in three submissions based on a number of factors including application of theory, strength of the research methods, and validity of conclusions reached. In well-organized conferences, anonymous written feedback from reviewers can be helpful in pointing out areas in need of improvement and to

reveal hidden assumptions overlooked by the researchers. For those papers that are accepted for presentation, further feedback can be provided verbally by conference attendees present when the paper is presented.

Research need not go through the conference step but might instead move directly into the publication stage of the theory-building process, which, again, opens the work to professional scrutiny. Publication widens the circle of scholars and practitioners available to critique and perhaps use the written work to further extend theoretical principles put forward. Publication requires further peer review to ensure the validity of the methodology and theoretical principles presented in the work and may take the form of books, book chapters, or journal articles.

Although it violates the spirit of objective scientific research, status of the publication plays a strong part in whether a proposed new theory gains traction in the academic world. *Top* journals and book publishers have greater name recognition and respect among scholars and the public press than lesser known publishers and thus command more attention. This is particularly important regarding the public press which, as an industry, has much more leverage to promote or discredit ideas in the global market than do academic presses.

As hypotheses and research questions garner attention and develop growing bodies of predictive and explanatory knowledge within the community of scholars and practitioners, the cases to which tentative theories can be applied widens. At the same time, alternative explanations also might be developed to test the validity and scope of initial hypotheses.

Once proposed, theories are continually tested, not to prove them true as is the popular conception but, rather, to prove them false. Through their assertions regarding predictable, observable cause-effect relationships, theories suggest where researchers look to test the theory. By trying to find conditions and circumstances in which theoretical predictions are not present or accurate, researchers can suggest theoretical limitations and perhaps rejection of theories in part or in whole. On the other hand, research that fails to disprove theoretical assumptions or that supports theoretical predictions in previously unexplored

circumstances, in turn, strengthens the theory by demonstrating its validity and reliability.

Besides the option of testing theories empirically, theories can also be evaluated at face value by examining a number of constitutional factors. Such evaluation determines whether new theories are accepted, tested, used, and further developed.

A number of criteria are used to judge the strength and value of theories. Theories should be clear in their description of what concepts are involved and how they are defined, the extent of phenomena covered, their explanatory or predictive purpose, and their succinctness. Theories also should be concisely worded, consistent, coherent, and avoid unnecessary conceptual and verbal complexities. They should make significant contributions to their respective fields—and ideally, to other fields of study as well. Theories should explicate their own contextual limits while leaving latitude for further application and expansion. Finally, as Whetten notes, theories that are both comprehensive and useful in their breadth and depth of application will have more chance of being accepted than those with narrow ranges of application.

Over time, what had once been a tentative assertion develops enough scrutiny, support, validation, and potential explanatory power to be considered a widely recognized theory. This process could take years and might never reach this stage if the theory is disproven along the way. Conversely, some theories accrue lasting power to the extent that they are repetitively supported in a wide range of times and places.

Popular conception is that given enough verification and scrutiny, a theory becomes a law or a fact. This line of reasoning is erroneous, owing to a misunderstanding of what a theory is. Whereas a theory is an explanation of why and how things occur, facts are verifiable occurrences. Therefore, it is safe to say that the function of theories is to explain and predict facts. As long as the explanation holds true for verifiable facts, the theory is considered valid.

This is not to say, however, that facts cannot be disproven. It was once considered a fact that a specialist could determine personality traits from the shape of a person's skull. This supposed fact was informed by the theory of phrenology, but both theory and *fact* have long since been

discredited. On the other hand, facts can remain valid even when theories devised to explain them have been disproven, thus opening the way for new theory generation.

Similarly, theories do not evolve into laws over time, contrary to popular perception. Laws describe relationships and effects, whereas theories explain how and why those relationships occur. Whereas laws reliably predict *what* will happen, theories explain *why* they happen. Once established, laws are far less malleable than are theories, which are continually refined to encompass a wider range of facts. Therefore, it would be more appropriate to say that theories can expand to include laws, but they do not morph into laws.

Under some circumstances, it is possible for researchers to become protective of theories, preventing scrutiny necessary for the theory's vitality and growth. As Robert Wyer points out, researchers can be bound by an over-dedication to the theory with which they have become accustomed and might become professionally associated with. In such cases, well-meaning theorists can become hampered by the conceptual boundaries defining the theory in question. The danger in such cases is that theories can become dogmatic. Rather than looking for weaknesses in theory, researchers can potentially become so committed to a theory that their research becomes a repetitive series of tropes to illustrate and support the theory, with little opportunity for theoretical expansion or for critique.

Salience of Theory Construction

Overt attention to theory construction waxes and wanes through time and across disciplines. Evident interest in the theory construction process is often a reflection of external sociopolitical events, on one hand, and ongoing debates within academe surrounding the need for theory, on the other. For example, continued social awareness of growing economic disparities could lead to renewed interest in theories inspired by Marxist philosophy as ways to explain those disparities. Also, as fields grow in terms of professional interest and research, the need for theory to explain bases of the field may arise in the effort to move the field beyond the initial stages of practical application of a limited set of concepts.

Interest in new theory construction can wax and wane in conjunction with cross-paradigmatic conflict. Academic friction between positivist and constructivist paradigms is a case in point wherein the nature and purposes of theory are continually debated. Indeed, this particular conflict has, since the early 1980s, raised fundamental questions about how theories are constructed, asking whether theories can be truly rational and separate from negotiation and compromise conditioned by historical circumstances.

Whether or not it is a predominant topic of research interest in itself or a practical means of explaining reality, theory construction is a crucial factor in the ways that we understand and engage with observable phenomena within all fields of the humanities and sciences.

Warren J. Bareiss

See also Hypothesis Testing; Scientific Revolutions; Theory Across Disciplines; Theory Change; Theory Structure

Further Readings

- Cornellissen, J. P. (2006). Making sense of theory construction: Metaphor and disciplined imagination. *Organization Studies*, 27(11), 1579–1957. <https://doi.org/10.1177/0170840606068333>
- DeFleur, M. L. (2010). *Mass communication theories: Explaining origins, processes, and effects*. Pearson Education.
- De Jong, H. L. (2010). From theory construction to deconstruction: The many modalities of theorizing in psychology. *Theory & Psychology*, 20(6), 745–763.
- Friedman, K. (2003). Theory construction in design research: Approaches and methods. *Design Studies*, 24(6), 507–522. [https://doi.org/10.1016/S0142-694X\(03\)00039-5](https://doi.org/10.1016/S0142-694X(03)00039-5)
- Glaser, B. G., & Strauss, A. L. (1967). *The discovery of grounded theory*. Aldine.
- Haig, B. D. (2008). An abductive perspective on theory construction. *Journal of Theory Construction & Testing*, 12(1), 7–10.
- Kuhn, T. S. (1966). *The structure of scientific revolutions*. University of Chicago Press.
- Rivard, S. (2014). The ions of theory construction. *MIS Quarterly*, 38(2), iii–xiii.
- Timmermans, S., & Tavory, I. (2012). Theory construction in qualitative research: From grounded theory to abductive logic. *Sociological Theory*, 30(3), 167–186. <https://doi.org/10.1177%2F0735275112457914>
- Weick, K. E. (1989). Theory construction as disciplined imagination. *Academy of Management Review*, 14(4). <https://doi.org/10.5465/amr.1989.4308376>
- Whetten, D. A. (2002). Modeling-as-theorizing: A systematic methodology for theory development. In D. Partington (Ed.), *Essential skills for management research*, 46–71. Sage Publications.
- Wyer, R. S., Jr. (2004). A personalized theory of theory construction. *Personality and Social Psychology Review*, 8(2), 201–209. https://psycnet.apa.org/doi/10.1207/s15327957pspr0802_14
- Zhao, S. (1996). The beginning of the end or the end of the beginning? The theory construction movement revisited. *Sociological Forum*, 11, 305–318.

THEORY STRUCTURE

Across the sciences, practitioners use theory to explain and predict a wide range of phenomena. Some type of theory is found in virtually all scholarly disciplines and areas of inquiry—in the physical, life, and natural sciences; the social sciences; and in the arts and humanities. As a result, theory—and what constitutes a good one—means different things to different people. For example, for those in the physical sciences, theories that address interacting or colliding elements (e.g., particles) seek to explain a change in state of one or more of the elements. In contrast, theories in the social sciences are often more abstract than physical theories; they tend to lack a mechanism that is thought to impel the elements, specified as variables, to change in magnitude on a single attribute. For example, an individual's (i.e., an element's) contact with dissimilar others (an attribute) is thought to increase the individual's tolerance of dissimilar others (another attribute). The causal relationship is present, but the mechanism is not.

Such differences in theoretical orientation lends to theories being classified as either scientific or nonscientific, with those theories that logically predict observable patterns of phenomena (scientific theories) in one camp standing in

distinction from those with a more ideological and/or informal nature (nonscientific theories). For the purposes of developing objective knowledge about the way the world works, scientific theory is the most preferred across disciplines. Scientific theories feature a structure that one can order and delimit. The most prevalent formal classifications that represent structures of scientific theory are *theories as axioms* and *theories as models*, which are closely coupled with the *syntactic view* and the *semantic view* of theory. The following sections discuss such distinctions, summarizing how theory is structured across the sciences.

Theories as Axioms

Throughout the history of science, axiomatic theory—often referred to as *deductive theory*—has been the accepted framework/structure of scientific inquiry. Philosophers of science have termed this mode of theory the *received view*, due to its widespread use across disciplines until the 1970s. Central to axiomatic/deductive theory is that all predictions and propositions stem from axioms, otherwise known as *postulates*; it is from such statements that theories originate. Axioms refer to ideas or knowledge claims we take for granted—ideas that are assumed to be true in order to understand that which follows from them. Axioms contain abstract, interrelated concepts—semantic devices that denote phenomena—that state the essence or process of something in the world. Further, axioms are generally not provable (which is why they are axioms/postulates rather than facts or laws). In physics, an axiomatic statement might be that “space is three-dimensional.” In biology, an axiom might take the form of “cells are the basic unit of life.” A general axiom in science is that the “universe is rational,” or put differently, x is not equal to $-x$.

Axioms exhibit a hallmark of deductive theory—a syntactical structure. Theories of this nature are illustrated by sentences consisting of both a logical and extralogical vocabulary. Logical symbols in theory have their foundation in mathematics and include connectives such as *and*, punctuation symbols such as brackets ([]), variables (such as x , y , and z), and symbols representing varying states

of equality ($=$, $\neq$). Those who are unversed in formal logic are likely more familiar with theories stated using extralogical vocabulary. Extralogical vocabulary consists of a theory’s descriptive terms, or concepts, which refer to observable or unobservable phenomena. Concepts rely on their definitions for understanding in the sense that one is only given meaning when defined. For example, one definition of a *role* is a sequence of gestures marking a course of action. This allows investigators of all types (students, researchers, and philosophers of science) to comprehend the phenomenon denoted by the concept. Ideally, the concept will take on a uniform meaning such that all investigators are in agreement when dealing with the phenomenon.

In the structure of theory, concepts are distinguished by whether they simply label a phenomenon or refer to one that pertain to differences in magnitude. Concepts of the first kind include the atom and electron; concepts of the second kind include atomic mass and electrical charge. Concepts that simply label phenomena are abstract and do not refer to concrete times and locations—they can be generalized. In contrast, concepts that relate to magnitude may vary in terms of size, weight, speed, and force. When we define concepts by their variable properties we have, *ipso facto*, translated them into variables. Thus, scientific theories often seek to demonstrate how variation in one phenomenon explains variation in another.

According to the received view, concepts are connected to one another via interrelated statements. When employed correctly, the statements follow a process of deduction, beginning with axioms/postulates and then deriving propositions from them. Until recently, this was the theoretical structure most often used in the sciences and the one most advocated in philosophy. At its fundamental level, a proposition should specify the connection between two or more variables. For example, the proposition “the stronger the degree of social integration in a society, the lower the society’s rate of suicide” contains two variables (*degree of social integration* and *suicide rate*). The propositional statement not only connects these two variables but also asserts an indirect and inverse relationship. Propositions can also differ in their level of abstraction, which is demonstrated

frequently in the social sciences. As such, a series of propositions may seek to explain all economic crises in general while another propositional system may link empirical facts concerning the Great Depression. This distinction leads to a common question about theory: Are empirical formats, which consist of generalizations that do not stem from axioms, really theoretical? One supported empirical generalization is that democratic nations, rarely, if ever, go to war with one another. This proposition, from the discipline of political science, is not very abstract, and deals with the observable concepts *democracy* and *war*. Furthermore, the concepts do not characterize a whole class of phenomena. If this were the case, *democracy* would be replaced by *political systems* and *war* with *conflict*. Ideally, relationships in a theoretical system should be deduced from a highly abstract assumption (axiom), but this is not the case. Instead, linked as it is to empirical contexts, it is simply a generalization that needs a theory to explain it. Even so, scholars in many areas would not hesitate to declare such statements as a *theory*.

To be sure, a rigid, parsimonious, and robust axiomatic/deductive theory is difficult to achieve. In fact, among philosophers of science, there is little consensus on which *theories* precisely adhere to the principles of pure, axiomatic theory. Nevertheless, when structured appropriately, there is arguably no theoretical form more powerful. Given the interrelatedness of axioms and propositions, empirical support for one or more of the propositions simultaneously signals support for the axioms from which they were derived. This is more plausible in the physical and biological sciences, where concepts are stated with more precision and exogenous factors are more controlled for than in the social sciences. Even so, the structure has fallen out of favor from scholars of all stripes during the last few decades as alternative, (perhaps) more practical approaches have emerged.

Theories as Models

The structure of what is accepted as theory changed from the *theories as axioms* perspective to the *theories as models* perspective roughly a half century ago. This shift occurred for various

reasons including the notion that treating theories as models allows for more flexibility in theory construction and applicability to a wider array of phenomena. An irony of this shift is that while many abandoned the received view for the *improved* models view, there is no consensus on what actually constitutes a *model*, or how models operate as theory. As summarized by Carl Craver, models are structures that support the validity of a theory, whereby (1) theories specify or define abstract, idealized systems; (2) models are the structures that satisfy these specifications or definitions; and (3) models are similar to some phenomenon (or system) found in the real world. Models thus are useful tools for understanding and predicting real-world events/processes, assuming a model indeed is an accurate proxy of that which it represents. In this view, models can be representations or simulations (e.g., a scaled abstraction of something, such as an atlas, map, or globe), abstract logical statements (e.g., mathematical expressions that stand for an empirical process), analogues (e.g., using Easter Island's societal collapse due to overpopulation as a comparative model for understanding potential outcomes of contemporary societies with population pressures), or experiments that serve as proxies for understanding real-world phenomena (e.g., the experiments by psychologist Stanley Milgram on obedience to authority).

Jonathan Turner conceives models as constructive tools that represent a range of activities; models involve concepts and their representations as a picture—an image, so to speak, of the crucial elements of a system or process. In this perspective, models are analytic devices that relate concepts to one another—they are diagrammatic representations of phenomena or events. Models as diagrammatic representations include (1) concepts that denote and isolate specific elements or features in the world, (2) the purposive arrangement of these concepts into a systematic imaginary or visual space so that they signify their interconnectedness, and (3) some symbolic marker that reveals the meaning of such interconnections among the elements in the model. Following this, models take two general forms: one *analytical* and one *causal*. Analytic models are more abstract, addressing generic properties of the universe. Analytic models depict

a complex set of connections among concepts/variables in an explanatory system. Analytic models are appropriate for connecting concepts described in an axiomatic or formal theory. Causal models are less abstract and more empirical, often addressing the individual or particular properties of empirical cases. Causal models are appropriate for understanding how measured concepts (variables) relate to each other and how such concepts account for variation in a specific outcome. Causal models often—but not always—represent the direct causal connections among elements in a system. For example, a causal model might represent the effect of population size in a society on the differentiation in that society's division of labor as a linear (or curvilinear) function. By constructing a model based on these criteria, one can identify the causal patterns that exist among a wide array of concepts/variables that exist in a system. Models thus provide scientists with a relatively simple, elegant mechanism by which to understand process, causality, and change in phenomena across scientific disciplines.

Significant differences between the theories-as-axioms and theories-as-models view are often difficult to identify. The most obvious difference is that the theories-as-models perspective moves away from propositional deductive statements and toward abstract representations or schemas that depict some phenomena. Other differences regard the fact that the theories-as-axioms view is defined by *syntax*; because of this axiomatic theory is thus sometimes referred to as the *statement view*. The theories-as-models view is more defined by *semantics*; because of this it is sometimes called the *nonstatement view* (see Craver's useful discussion of these distinctions).

Debate over a theory's syntactical and semantic components is crucial given that the failings of either can weaken all results stemming from a theory's research program. The debate, which has continued since the 1960s, concerns whether either view can adequately describe scientific theories. Proponents of the syntactic view stress its formal nature and connections to experiments and observations. Sentences are written in *higher order logic*, with one set of terms reserved for observational language. Critics of this approach

feel that formalization in such logic is cumbersome and that theorists should rather describe observational data as occurring in a physical space composed of varying inputs and outputs. This alternative is articulated in the semantic view, which relies more on the adaptability of mathematics than the rigidity of logic. To its critics, the semantic view is too vague when one must interpret what precisely a theory is explaining (see James Mattingly's critique of the semantic view for an elaboration on this theme). Further problems borne out of this debate involve how new theories come about from seemingly problematic older ones. As a result, hybrid approaches including tenets of both the syntactic and semantic views have been proposed.

Additional Dimensions of Theory Structure

In addition to the axiomatic and models perspectives described above, the structure (or components) of a theory can be classified along other dimensions. One such dimension regards whether a theory is *formal* or *informal*. Formal theories often have deductive propositions or interrelated statements that follow from one another and often involve logical or mathematical expressions. An example of a formal theory is Guillermina Jasso's theory of comparison processes, which uses formal expressions to represent how individuals compare resources to one another: the axiom of comparison states that there exists a Class Z of human individual phenomena that are produced by the comparison of an actual holding (A) of a good (or a bad) to a comparison holding (C) of that good (or a bad), such that Z is an increasing (or decreasing) function of A and a decreasing (or increasing) function of C :

$$Z^G = Z^G(A, C) \quad \partial Z^G / \partial A > 0 \quad \partial Z^G / \partial C < 0$$

$$Z^B = Z^B(A, C) \quad \partial Z^B / \partial A < 0 \quad \partial Z^B / \partial C > 0$$

By contrast, informal theories are less rigid conceptual frameworks that describe some phenomenon or process, often using narratives to explain the nature of something. Standpoint theory, a controversial theoretical framework that has

emerged in sociology, provides an example of an informal theory.

Another dimension of theory structure regards the *scope conditions* of a theory or the range of phenomena a theory purports to describe or explain. For example, in social psychology, expectation states theory (EST) is a predictive theory of behavior that examines the influence of status on individuals' performance in groups. EST has two scope conditions in which the theory is predictive: in task orientations (when group members are motivated to solve a problem) and in collective orientations (when group members consider it legitimate and necessary to consider each other's contributions). Scope conditions are important in theories as they delimit the explanatory range of a predicted outcome.

Another dimension of theory structure regards *level of analysis*. Theories—particularly in the social sciences but not unique to them—can be conceived as explaining micro-, meso-, or macro-level phenomena, depending on the area of inquiry. For example, theories addressing processes at the atomic level might be considered microlevel theories (such as Dalton's atomic theory); theories examining the nature of the larger universe would be considered macrolevel theories (such as Doroshkevich, Sunyaev, and Zel'dovich's theory of galaxy formation from cluster pancakes). There is no ultimate bifurcation point or ultimate consensus on what would define a theory at any level of analysis, and such distinctions are usually only relevant for placing a theory within an understood domain or common area of inquiry. Levels of analysis can be important though as some theories predict outcomes at one level of analysis differently than for others (e.g., processes at the quantum [micro] level are conceived to operate uniquely as compared to processes at other levels of analysis).

In general, theories consist of components—differing structures—that help define, explain, and predict phenomena. Although what constitutes a theory's structure may differ in form and the level of abstraction from theory to theory, all theories—regardless of structural form and complexity—are aimed at assisting scientists and scholars in better understanding the world.

Michael J. Carter and Christopher Lawrence

See also Axiom Schema; Axiomatic Theory; Concept; Fact Versus Theory; Framework; Modeling; Theoretical Vocabulary; Theory Construction; Theories, Semantic Conception of; Theories, Syntactic Conception of

Further Readings

- Craver, C. F. (2002). Structures of scientific theories. In P. Machamer & M. Silberstein (Eds.), *The Blackwell guide to the philosophy of science* (pp. 55–79). Blackwell.
- Gonzalez, W. J. (2014). *Bas van Fraassen's approach to representation and models in science*. Springer.
- Jasso, G. (2002). Formal theory. In J. Turner (Ed.), *Handbook of sociological theory* (pp. 37–68). Kluwer Academic/Plenum.
- Mattingly, J. (2005). The structure of scientific theory change: Models versus privileged formulations. *Philosophy of Science*, 72(2), 365–389. <https://doi.org/10.1086/432428>
- Suppe, F. (1974). The search for philosophic understanding of scientific theories. In F. Suppe (Ed.), *The structure of scientific theories* (pp. 266–307). University of Illinois Press.
- Suppe, F. (1979). Theory structure. In P. Asquith & H. Kyburh (Eds.), *Current research in philosophy of science* (pp. 317–338). Philosophy of Science Association.
- Turner, J. H. (2003). *The structure of sociological theory* (7th ed.). Wadsworth.

THEORIES, SEMANTIC CONCEPTION OF

The semantic conception of theories holds that a scientific theory can be identified with a collection of models. This view was originally put forward by Patrick Suppes in “A Comparison of the Meaning and Uses of Models in Mathematics and the Empirical Sciences” with the aim of repudiating the *received view* of theories then popular in logical positivism. This entry discusses the foundations and development of the semantic conception of theories, starting with the view that theories should be viewed in terms of models. We will also look at the advantages of a model-based approach, and how models could provide a better pathway to discovering the laws of nature.

The Classical View of Theories

In the early days of symbolic logic, theories tended to be *axiomatized*—for example, arithmetic, as illustrated by Peano's axioms, and the geometry of Euclid, as formalized by David Hilbert (1862–1943).

These mathematical theories involve a kind of axiomatization with some of the axioms thought of as definitions and others as laws. This classical view is associated with the *H-D view* of confirmation as used by Carl Gustav Hempel (1905–1997), the idea being that you confirm a theory by using it to make predictions about an observable phenomenon, then check to see whether the predictions are correct. For Hempel, predictions are structurally identical with explanations: The premises need to include at least one law and some initial conditions, then a statement is deduced about something observable. Put simply, the classical view looks at theories as collections of sentences. So the classical logical positivist view construes a theory as a set of sentences in an axiomatized system of first-order logic.

Foundation of the Semantic Conception of Theories

Bas van Fraassen (b. 1941), a key founder of the semantic conception of theories, has strongly criticized the syntactic view for focusing attention on technical questions which are irrelevant from a philosophical point of view. The discussions of axiomatizability that take place in restricted vocabularies and use *theoretical terms* lead to unnecessary self-generated problems. Rather than lend itself to an axiomatic approach, science, according to van Fraassen, should be seen as defining mathematical models directly, with the focus on the connections between theoretical terms and language that are more readily understood.

During the 1960s, philosophers of science began to explore different approaches to scientific theories. One prominent approach was the semantic conception, according to which theories are viewed as collections of models together with hypotheses describing how these models relate to different areas of nature. There were numerous

views about how to characterize these models, and whether they should be looked at as abstract mathematical structures, open to precise formal specification, or whether they should be taken to be more concrete, like the molecular models built by chemists.

The semantic view repudiates the syntactic view of theories, upheld by the logical positivists and empiricists. The *syntactic conception* of a theory affiliates it with a body of theorems, in a language particularly selected for that theory. The alternative semantic conception sees the theory as being identified with a class of structures as its models in language which is neither basic nor unique—that is, the same class of structures could be described in quite different ways, within its own limits, thus ensuring the models take center stage. Hempel and Rudolf Carnap (1891–1970) are two main proponents of the first, syntactic view of theories.

The semantic view holds that theories as models or sets of models. Some semantic views suggest that theories should be identified with a Tarskian class of set-theoretic models, while others propose that models should be set out in mathematical language that relates to the field in which the theory belongs as a member. Whereas semantic views differ, they all take as central to their notion the model as the unit of scientific theorizing. (Polish logician Alfred Tarski [1901–1983] undertook work that would lead to rigorous definitions for notions useful in science methodology. He published a paper that would give his analysis of the notion of a true sentence for a given formal language. For him a language is fully interpreted if all its sentences have meanings that make them true or false.)

A *model* is a system of semantic rules that interpret an abstract calculus, thereby applying scrutiny to the semantics of a specific language. For example, in the kinetic theory of gases, if we take the mathematics used in the theory and reinterpret the terms of calculus so that they refer to billiard balls, the billiard balls are a model of the kinetic theory of gases, a semantic model. Logical positivists, such as Hempel and Carnap, believed that such models are superfluous to scientific study, holding that they are at most only of pedagogical, psychological, or aesthetic use. Proponents of the

semantic view of concepts, such as Franz von Suppé (1819–1895), van Fraassen, and Ronald Giere (1938–2020), think the opposite and that we should reject formal calculus and take the view that a theory is a group of models.

Critics of the semantic view believe that theories are not constructed in this way; models are generally independent from theory, rather than being part of them. They are in many ways autonomous agents. This can give them two aspects: *construction* and *functioning*.

Construction of Theoretic Models

When we look at how models are constructed in actual science, we can see that they are derived completely neither from data nor from theory. Theories do not provide the algorithms for constructing a model; the process is not akin to a vending machine—we cannot put problems in and receive a ready-made model. A good example of this is the London model of superconductivity. There is no justification for the model's foremost equation (it cannot be derived from electromagnetic or any other theory). It is motivated purely on the basis of phenomenological pretexts. Basically, the model has been constructed from the *bottom up* rather than from the *top down*, giving it a good deal of independence from theory.

That models are independent also means that they perform functions they could not perform if they were part of, or strongly reliant on, theories.

To recap: According to the semantic view, a theory is a model or set of models.

The model, or models, may be accompanied by theoretical hypotheses that link the model with some real-world phenomenon. A model can be large scale, a map, a diagram, or an analogy. More commonly in science, we have mathematical models: Sets of equations that are thought to capture real-world phenomena in an idealized form.

The Role of Semantic Models in Science

Increasingly in the neurosciences and in psychology, there is a need for computational models or computational architectures, or programs devised

to capture some aspect of the human condition. Semantic models can fill the gaps in theories like classical mechanics and quantum physics, theories that do not represent anything in the material world of (naive) realism. They can provide a concrete representation filled with the details of a specific situation, rather than a theory standing on its own.

Models can achieve results where theories cannot, for example, when theories are too complicated to handle. A simplified model can provide a solution. Quantum chromodynamics, for instance, is a theory that is difficult to use to study the hadron structure of a nucleus, although the theory is fundamental in order to solve the problem. One way around this was to construct a model to describe the various degrees of freedom of the system being considered. The advantage here is in getting actual results rather than silence.

The disadvantage is that it is not at all easy to understand the relationship between theory and models, since the two are in a sense, contradictory. In a more extreme case, we can use a model to help where there are no theories available at all. This is particularly common in the fields of economics and biology, where there are sometimes no overarching theories to call on. In this case, scientists tend to work with *substitute models*.

Relatedly, we can see a theoretical model as a kind of definition. In particular a theoretical model defines a certain kind of system, for example, a classical particle system or, in genetics, a Mendelian system.

Advantages of Using a Semantic View

The benefits of using a semantic conception of theories include the fact that because the theory is based on a model it cannot be false, since it is just a big definition of a *kind* of system. Thus, we do not need to figure out which parts of the theory are definitional and which are empirical, which could be a positive factor. With regard to Newton's second law, $F = ma$, is this an empirical claim or a definition of force? Or is it a partial definition of mass or acceleration? It is not clear.

Viewing a theoretical model as a definition, however, is not to say it does not have empirical

content. Content is included by way of the theoretical hypotheses. The theoretical hypothesis supplies the hypothesis that a real physical system is a system of the type defined by a model. We could put this as saying that a theoretical hypothesis says that some part of the real world is similar to the model.

This brings us to issues of realism. Van Fraassen believes that realism sees theories as literally true. Whether this means exactly or completely true is not apparent, but on this basis it suggests that a model is never either exactly or completely true: The resemblance between model and reality can never be absolutely perfect. But when it comes to some things, we should not interpret realism in such a way as to require it to be strong. It is sometimes claimed that the practice of building models favors anti-realism over realism.

Anti-realists point out, however, that the truth is not the main goal of scientific modeling. Nancy Cartwright puts forward several case studies that illustrate that good models are often false, and that supposedly true theories may not help when it comes to understanding, for instance, the working of a laser. Realists reject the view that the falsity of models makes a realist approach to science impossible, believing that a good model, even if it is not actually true, has a close approximation to truth.

Ronald Laymon argues that a model's predictions become better as we relax the idealization of the world, a view which he purports to support realism. The notion of an approximate truth is, however, too vague for anti-realists. There is no reason to assume that a model can be improved by de-idealizing it even if this were in line with accepted scientific practice, which it is not. Rather than putting effort into repeatedly de-idealizing existing models, they might better be shifting to an entirely new modeling framework once the adjustments become highly involved. A further problem with the concept of de-idealization is that it is not *controlled*. It is not clear how one could de-idealize a model to arrive at the supposedly correct underlying theory.

Laws of Nature

Science aims to discover the laws of nature, and philosophy has been given the challenge of explicating what these laws are. The two dominant accounts are the *best systems* approach and the *universals* approach, and these propose that the laws of nature apply to everything in the world.

This position does not appear to square with a view that imagines models at the center of scientific theorizing. What role can general laws play in science if models represent what is happening in the world, and what is the relationship between models and laws? One response could be that the laws of nature govern processes and entities in a model, rather than those in the world. In this scenario, fundamental laws do not state facts about the world but are true of the entities and processes in the model.

With no real counterpoint to this view given by realists, the challenge remains to understand how and if a realistic understanding of laws and a model-based approach can be reconciled.

Semantic views have many credible features, for example, they provide a way to discuss aspects of science without recourse to a particular language. They also seem to do more justice to scientific fields that resist axiomatization. One good example of this is Charles Darwin's theory of evolution. When the axiomatic approach was in its heyday, philosophers attempted to show how the evolutionary theory could be brought within the confines of orthodox conception of theories.

What happened was that they ended up producing formal theories bordering on triviality and ridiculous questioning about whether the theory of evolution was a tautology. The semantic conception then looked like a far more sensible approach, since it shed useful light on theoretical issues arising in contemporary evolutionary biology. In *The Structure and Confirmation of Evolutionary Theory*, Elizabeth Lloyd argues that the semantic account is more appropriate and powerful.

The semantic conception suits the practice of idealization, an often neglected area of science. Rather than seeing scientists as desiring to give literally correct descriptions of objects in the world, the semantic conception proposes that they instead conceive models accompanied by

claims that particular parts of nature correspond to these models in specific respects and to specific degrees. The debate over scientific models has major repercussions in the world of the philosophy of science. Scientific realism, reductionism, explanation, and the laws of nature were couched in terms of theories because theories alone were believed strong enough to carry the weight of scientific knowledge.

Pablo Lorenzano

See also Axiomatic Theory; Realism in Science; Theories, Syntactic Conception of

Further Readings

- Carnap, R. (1938). Foundations of logic and mathematics. In O. Neurath, C. Morrison, & R. Carnap (Eds.), *International encyclopedia of unified science* (Vol. 1, pp. 139–213). University of Chicago Press.
- Cartwright, N. (1983). *How the laws of physics lie*. Oxford University Press.
- Giere, R. (1988). *Explaining science: A cognitive approach*. University of Chicago Press.
- Laymon, R. (1987). Using Scott domains to explain the notions approximate and idealized data. *Philosophy of Science*, 54, 194–221.
- Lloyd, E. (1994). *The structure and confirmation of evolutionary theory*. Princeton University Press.
- Suppe, F. (1989). *The semantic conception of theories and scientific realism*. University of Illinois Press.
- Suppes, P. (1960). A comparison of the meaning and uses of models in mathematics and the empirical sciences. *Synthese*, 12, 287–301.
- Suppes, P., Costa, N. C. A., & French, S. (1960). The model-theoretic approach in the philosophy of science. *Philosophy of Science*, 57, 248–265.
- van Fraassen, B. C. (1980). *The scientific image*. Clarendon.

the early 1970s for a variety of reasons. According to the positivists, a scientific theory consists of a deductively closed set of sentences in a vocabulary partitioned into an observational part and a non-observational, or theoretical part. The cognitive significance of the latter required special justification given the rejection of metaphysics as void of content on positivism's preferred account of meaning. According to that account, a statement can be regarded as either true or false only if either (1) it is analytic or contradictory, in which case it has logical significance, or (2) it can be tested directly by observation or experiment, in which case it has empirical significance. Metaphysics, it was claimed, has neither. But one did not want the same verdict for the sentences of scientific theories essentially involving theoretical vocabulary. Two programs resulted. One attempted to provide the theoretical language with indirect or partial empirical significance. The other aimed to show that theoretical vocabulary can be eliminated without detriment to the aim and function of scientific theories. Neither program developed satisfactorily, and the syntactic view was dropped for a variety of reasons.

Formative Influences

The framework out of which the syntactic conception grew was laid by logical positivists of the Vienna Circle during the 1920s and early 1930s, especially with the verifiability theory of meaning. The development of the syntactic view began with the publication of "Testability and Meaning" in two parts in 1936–1937, only after its principal founder, Rudolf Carnap, had emigrated to Chicago.

Logical positivists drew inspiration from two main influences. One was the development of formal logic in the late 19th and early 20th centuries, which considerably aided progress in the foundations of mathematics by using formal languages and the axiomatic method. Mathematical theories so formalized included arithmetic (Giuseppe Peano), Euclidean geometry (David Hilbert), and Cantorian set theory (Ernst Zermelo). Among the fundamental unsolved problems in mathematics, Hilbert included, as the sixth, a mathematical axiomatization of physics.

THEORIES, SYNTACTIC CONCEPTION OF

The syntactic conception of theories, often called the *received view*, originated with logical positivism in the 1930s and began to fall out of favor in

The other driving factor derived from an antipathy to the philosophical idealisms of G. W. F. Hegel and Francis Herbert Bradley and to metaphysics in general. This led to the verificationist account of meaning. The meaning of a statement is the method of its verification: no method, no meaning. An example that impressed Moritz Schlick, the founder of the Vienna Circle, was Albert Einstein's 1905 critique of absolute simultaneity. The claim that Event A is simultaneous with Event B has no meaning in the absence of a means of verification. There is no frame-independent technique. Rather one has, as introduced by Einstein, a frame-relative method by using light signals to synchronizing clocks at rest in the frame. Metaphysical claims are especially wanting. It eventually became clear, however, that the verificationist criterion is too strong. For example, universal claims cannot be conclusively verified since one cannot preclude the discovery of a counterexample. Hence, Carnap replaced the verificationist criterion with the weaker requirement of testability—that is, verifiability or falsifiability, and, in general, confirmation (or disconfirmation) to a degree. Metaphysics remains meaningless under this relaxation.

The Empirical Significance of Theoretical Terms

How, then, do the theoretical terms of science acquire meaning, or empirical significance? No problem arises if a theoretical term is explicitly definable in terms of the observation language. Suppose the term in question is a monadic predicate Px . It is explicitly definable in this way just in case there is an observation language formula $\phi(x)$, in which just the variable x is free, such that the sentence $\forall x(Px \leftrightarrow \phi(x))$ is a consequence of the theory. (Function terms, including individual constants, are handled similarly if certain existence conditions are fulfilled.) Then, if c is a name for an observable entity or else a previously defined theoretical term, the truth of the sentence Pc can be tested by testing whether or not $\phi(c)$.

Unfortunately, few theoretical predicates in actual theories are explicitly definable. To fill this void, Carnap introduced the notion of *reduction sentences*. To simplify, such a pair amounts to

specifying an observable condition sufficient for a theoretical claim ψ and an observable condition sufficient for the negation of ψ . This yields a decision procedure for ψ if it can be guaranteed that one or the other condition holds. But in the absence of both, whether ψ is the case remains undetermined.

Later positivists took a looser attitude, requiring only that there be so-called bridge laws among the axioms connecting theoretical with observation terms just in virtue of essential co-occurrence in the law. Via the bridge laws, the theoretical terms acquire (at least partial) empirical significance by *osmosis* from the observational roots of the theory.

The Eliminability of Theoretical Terms

Alternatively, one can seek to eliminate the theoretical language in its entirety and claim that the theoretical sentences together with bridge laws amount to *inference tickets* between observation sentences. This stance is often referred to as *instrumentalism*.

The Theoretician's Dilemma

But how to eliminate the theoretical language? Explicit definability of all theoretical terms would suffice, but we have seen that this is unrealistic. Another method is to presume that the theory is formulated in first-order logic, so that the set of consequences of the theory is effectively enumerable. The set of all observation sentences is therefore effectively enumerable simply by deleting any sentences containing theoretical terms from the enumeration. Moreover, William Craig showed that any effectively enumerable set of sentences is axiomatizable. Thus, one recovers an axiomatized theory entirely in the observation language, having exactly the same observational consequences as the original theory.

Nonetheless, success here just raises the question of what purpose theoretical terms might serve beyond mere convenience. If none, the theoretician has the dilemma that theoretical terms are in principle dispensable. Even though he had no proof, Carl Gustav Hempel countered that even if a theory provides the same deductive

systematization as its Craigian replacement, it may provide augmented *inductive* systematization. Using Jaakko Hintikka's two-dimensional account of inductive inference, Raimo Tuomela claimed to vindicate the indispensability of theoretic terms for inductive tasks.

Ramsey Eliminability and Empirical Content

So far, we have given the impression that the empirical content of a theory consists in the set of consequences in the observation sublanguage. However, there is a second candidate definable in terms of the models of the theory. Suppose the observational vocabulary consists of terms $O_1, \dots, O_n$ and the theoretical vocabulary terms $T_1, \dots, T_m$. Suppose further that the theory is formulated in first-order logic. A model $\mathfrak{A}$ of the theory has the form $(|\mathfrak{A}|, O_1^{\mathfrak{A}}, \dots, O_n^{\mathfrak{A}}, T_1^{\mathfrak{A}}, \dots, T_m^{\mathfrak{A}})$, where $\mathfrak{A}$ is a non-empty set and each $O_i^{\mathfrak{A}}$ and each $T_j^{\mathfrak{A}}$ are either a relation or an operation on $|\mathfrak{A}|$, depending on whether the term in question is a predicate symbol or a function symbol. For such a model, let $\mathfrak{A}_0$ be the restriction $(|\mathfrak{A}|, O_1^{\mathfrak{A}}, \dots, O_n^{\mathfrak{A}})$ of $\mathfrak{A}$ to the observation language, alternatively called the *reduct* of $\mathfrak{A}$ to the observation language. Must it be the case that $\mathfrak{A}_0$ is a model of the set of consequences of the theory in the observation language? For some theories yes, but for others no. If every reduct of the theory is a model of the set of its observational consequences, then the theoretical language is said to be *Ramsey eliminable*. For a theory whose theoretical language is Ramsey eliminable, it makes no difference if we represent the empirical content of the theory by the set of consequences in the observation language or by the class of reducts of models of the theory. But if the theoretical language is not Ramsey eliminable, then the class of reducts of models of the theory is the preferred representation of the theory's empirical content. For this class represents the possible observable situations permitted by the theory, whereas the set of consequences in the observation language is satisfied in certain circumstances ruled out by the theory.

Thus, if the theoretical languages are typically Ramsey ineliminable, it might point to a positive function for them. Joseph Sneed, in *The Logical Structure of Mathematical Physics* (1971), conjectured that mass and force are not Ramsey eliminable from Newtonian mechanics. Robert Rynasiewicz (1981) disproved the conjecture. Herbert Simon and Guy Groen thought that Ramsey eliminability holds generally for all reasonable theories, where the criterion of reasonableness demands that the theory be finitely and irrevocably testable (FIT). But this requires that every substructure of a model of the theory also be a model or, equivalently, that only universal quantifiers appear in the axioms when cast in prenex normal form. This rules out theories with existential commitments, and so, given the prevalence of these, FITness appears to be an unreasonable requirement.

Partial Structures and Empirical Content

Bifurcating the language is intended to distinguish between those properties and operations that are observable and those that are not. But we might also ask about which of the individuals in the domain of discourse $|\mathfrak{A}|$ of a model $\mathfrak{A}$ are observable. The preceding discussion implicitly assumes that each and every individual is observable, an unrealistic assumption for theories that trade in things too small to see. If there is indeed a workable observable/unobservable distinction, it applies foremost to events, situations, or states of affairs, while an individual or property or operation is derivatively one or the other insofar as it is a constituent of the event or situation. Moreover, a property, such as roundness, can be observable in some instances but not in others, depending on the individual involved. Thus, the putative distinction simply cannot be reflected in language as the positivists imagined. Also, a given individual may be observable in some types of circumstances but not others. The net result is that, to represent the empirical content of a theory, one needs to resort to fragments of its models that are not themselves fully fledged models, even for the observation language, but are only partial models in which, for some individuals and properties, it remains

indeterminate whether the individual has the property or not. We cannot get at the empirical content of a theory simply by bifurcating its language.

The Demise of the Syntactic View

Even before Carnap and the Vienna Circle, influential criticisms had been leveled against the received view. One was the charge that observation is theory laden. Another was the rise of scientific realism bolstered by the new causal theory of reference extended to natural kind terms.

These factors told against the bifurcation of the vocabularies of theories. Scientific realism, however, by and large continued to use the framework of theories as a deductively closed sets of sentences. Earlier on, even this had come under challenge largely because first-order logic was deemed too straitjacketing. Indeed, at best trivial examples of theories from mathematical physics had been axiomatized in first-order logic. This occasioned abandonment of linguistic features altogether in the development of the so-called semantic view, which holds that theories are rather classes of models, where the sense of *model* is broader than in logic.

Robert Rynasiewicz

See also Theories, Semantic Conception of

Further Readings

- Carnap, R. (1936–1937). Testability and meaning. *Philosophy of Science*, 3(4), 419–471; 4(1), 1–40.
- Carnap, R. (1956). The methodological character of theoretical concepts. In H. Feigl & M. Scriven (Eds.), *Minnesota studies in the philosophy of science* (Vol. 1, pp. 38–75). University of Minnesota Press.
- Hempel, C. (1958). The theoretician's dilemma. In H. Feigl, M. Scriven, & G. Maxwell (Eds.), *Minnesota studies in the philosophy of science* (Vol. 2, pp. 37–98). University of Minnesota Press.
- Rynasiewicz, R. (1981). *Varieties of eliminability of theoretical terms and the empirical content of theories* [PhD dissertation]. University of Minnesota.
- Simon, H. A., & Groen, G. J. (1973). Ramsey eliminability and the testability of scientific theories.

The British Journal for the Philosophy of Science, 24(4), 367–380.

Tuomela, R. (1973). *Theoretical concepts*. Springer-Verlag.

THOUGHT EXPERIMENTS, SCIENTIFIC AND PHILOSOPHICAL

Thought experiments abound in philosophical literature and in many areas of scientific inquiry. Scholarly interest in thought experiments has increased over the past several decades. Whereas the role and status of thought experiments has been a point of focus for historians and philosophers of science for some time, philosophers have more recently devoted significant attention to this topic through books dedicated to it, themed special issues of journals on thought experiments, and many studies in journal articles and book chapters of this or that particular thought experiment.

Why has there been so much focus on thought experiments? There are many possible explanations for this increased attention, but three reasons are noteworthy. First, thought experiments, obvious from their name, seem to have something in common with real-world experiments. Determining the nature of that apparent similarity has drawn the attention of many scholars. Another reason is that thought experiments, especially those in the sciences, purport to inform us about the actual world, whether about what is actually the case or about what *could* or *could not* be the case. If thought experiments have this ability, many have wondered how merely *thinking* about some scenario can provide such insight. In other words, many thought experiments purport to provide knowledge of the world simply from one's armchair, but this appears to run counter to the thrust of much scientific inquiry. A third reason for the heightened interest in the status of thought experiments is related to the second: Thought experiments do not appear straightforwardly to sit well with contemporary versions of empiricism. As a result, many have worked to understand the epistemic import of thought experience on empiricist grounds. The entry first addresses

the nature and types of thought experiments. Next it discusses several prominent examples of scientific and philosophical thought experiments.

The Nature and Types of Thought Experiments

A thought experiment requires the reader to imagine some situation contrary to their past, current or even possible experiences. Thought experiments may appeal to times that the author admits never existed in the past, such as Thomas Hobbes's *natural state of humans*, or to scenarios even more abstract containing features like frictionless planes. In such imagined scenarios, the reader is often first instructed to examine carefully the features of the alternative setup. After this examination of the initial conditions, some change is usually introduced, such as the collision of one body with another body or the recognition that resources have become scarce.

Another common feature of thought experiments is that they are experiential in nature insofar as there is a transition through different stages of the experiment. Rather than merely imagining some static, counterfactual situation, a thought experiment provides the reader with a sort of mental experience. Thus, Charles Darwin asks readers to imagine wolves with differing levels of fitness competing with one another for prey as the number of prey shifts over time, Judith Jarvis Thomson illustrates the possibility of waking up and discovering a violinist who is connected to you and in need of life support, Thomas Hobbes imagines humans living without civil relationships and in constant fear of attacks from others as they face competition for resources, and in *Le Monde* René Descartes portrays a fictional world with its parts in the process of coming together. Just like experiences in the actual world and contrived experiments in a laboratory, thought experiments happen in stages like these and occur over a (fictional) period of time.

So far so good, but how could sitting in one's armchair and imagining scenarios like these—some very different from everyday experiences—promise to provide knowledge that will be usable in the actual world? Broadly speaking, there have

been two prominent approaches to understanding how merely the process of considering a thought experiment could do this: the *a priori* camp and the empiricist camp. After describing each of these prominent approaches, this entry will mention a third view that argues that thought experiments are empirical experiments in the mind. The *a priori* camp holds that some thought experiments transcend what it is possible to know from experience and as a result give those who work through them access to truths beyond experience, for example, the laws of nature. In a 2004 paper, James Robert Brown argues for this understanding of the nature of thought experiments by linking intuitions used in thought experiments to intuitions in mathematics. A consequence of the *a priori* account of thought experiments is that it implies that scientific inquiry is constituted by both experienced-based evidence, such as laboratory experiments, and *a priori* intuitions that transcend that experience. Against the *a priori* view, the empiricist camp tries to understand thought experiments as arguments that are not different in principle from an argument that one would make on the basis of experience or experience. For example, in a 2004 paper, John Norton argues against the *a priori* view by contending that thought experiments simply draw upon what is already known about the world. Sometimes a thought experiment may make explicit what was tacitly already known about the world, but a thought experiment adds nothing in addition to what experience teaches. On this empiricist understanding, thought experiments do nothing logically over and above what an empirically based argument would do; instead, thought experiments make those arguments more vivid than they would otherwise be if they were presented in argumentative form. An alternative view of thought experiments to these two suggests that thought experiments are empirical experiments. In a 1992 paper, David Gooding argues for such an approach by proposing that, rather than being empirical *arguments* that have simply been made more vivid, thought experiments are actual experiments conducted in mental laboratories. This approach highlights the experiential, narrative aspects of thought experiments draws attention to the fact

that many are often open to empirically based criticisms, just as laboratory experiments are.

Authors who use thought experiments in scientific and philosophical contexts do so with different goals in mind. It may be difficult or currently impossible to test some prediction of a theory, so a thought experiment may try to bridge that gap. Sometimes thought experiments seek to disprove the possibility of a claim. Other times they show the possibility of some claim given the initial conditions of the thought experiment. Still others seek to show the necessity of some feature given some broader theoretical component that is held. Despite these varied uses, and the many others that could be catalogued, some have tried to group thought experiments into broad kinds. For example, Sorensen's (1992) book offers one of the most comprehensive philosophical treatments of thought experiments and divides thought experiments into *necessity refuters* and *possibility refuters*. Alternatively, Brown's influential (2011/1991) treatment of thought experiments taxonomizes them into *constructive* and *destructive* thought experiments. These and the taxonomies that others have offered have been controversial, so another approach simply divides thought experiments into their many and various uses by particular authors. Rather than deciding this question of taxonomy, the remainder of this entry will examine several prominent examples of thought experiments, first in scientific contexts and next in philosophical ones.

Two Scientific Thought Experiments

Galileo Galilei and Falling Bodies

In *Discorsi*, a work from 1638, Galileo used the genre of dialogue to show with a thought experiment that Aristotle's claim that the speed of falling bodies differs according to their weight leads to a contradiction. The thought experiment arises when, during a discussion of Aristotle's arguments against the void, the character Simplicio notes an assumption that Aristotle makes related to the manner in which objects that differ in *heaviness* move in the same medium; the assumption is that such objects move at different rates of speed according to a *ratio as to their weights*. Simplicio takes this to imply, for example, that an object

that is 10 times as heavy as another will move ten times as fast.

After some debate between the characters Simplicio and Salviati about whether Aristotle ever actually tested this assumption by dropping objects with differing heaviness, Sagredo asserts that he himself *made the test* to see whether Aristotle's assumption would be upheld. Sagredo reports that the assumption did not hold up when he himself tested it. Salviati then interjects and denies the need for this or for any further empirical tests. Instead, he states that with a brief *demonstration* it is possible to prove that Aristotle's assumption is false.

The brief demonstration is a thought experiment. It begins with Salviati asking Simplicio if he agrees that the only way for the speed of a heavy falling body to increase or decrease is by the use of force or by some impediment. Simplicio agrees to this assumption, and Salviati further asks him whether he agrees that if two bodies with unequal speeds were connected to one another (to become a single, linked body) then in doing so the faster one would be slowed by the slower one and the slower one would be sped up by the faster one.

After Simplicio agrees to this consequence, Salviati asks him to imagine two stones, one a large stone that is moved by 8° of speed and the other a smaller stone moved by 4° of speed. Given the assumptions that Simplicio has agreed to already, it follows that if these two stones were joined together the speed of the compound stone would be less than 8° of speed. However, the new compound stone formed from the previously separated stones is *heavier* than the original stone that was moved at 8° of speed, and so we would expect it to move at a greater rate of speed than the individual larger stone. The contradiction should now be clear: the compound stone should both move *more* swiftly and *less* swiftly than the original stone, which moved at 8° of speed.

The beauty of this thought experiment is its simplicity as the reader follows through its stages from the initial assumptions to the exposure of the contradiction. The contradiction that the experiment exposes leads the reader to reject the Aristotelian assumption, and as Salviati states this seems to be something that can be done without conducting any empirical tests.

Charles Darwin's Origin and the Illustrations That Darwin Provides

Darwin's *On the Origin of Species* (1859) collected together observational data from many seemingly disparate sources. Darwin's own thinking on speciation was impacted by many sources, such as his own field observations and notes he took about them, the observations and actions of domestic animal breeders, and the many reports that he read on topics that interested him, such as those on the fossil record. The *Origin* begins with four chapters that seek to outline the theory of natural selection, beginning with consideration of the way that variation occurs among domesticated animals. Near the end of Chapter 4 (1859, 90 ff.), Darwin offers what he calls *illustrations* of how natural selection acts. Rather than obviously relying upon any particular person's empirical observations, these illustrations are thought experiments, which may come as a surprise to some of Darwin's readers.

In one such *illustration*, Darwin asks the reader to imagine a wolf that preys on different animals and also that that wolf's fastest prey has increased in numbers. Alternatively, he notes that one could imagine the same wolf and that its slower prey had decreased in numbers so that there would be a greater ratio of fast to slow prey in the resulting available pool. Given such an initial setup, Darwin suggests that the fastest and *slimmest* wolves would have the best chance of survival and would, as a result, be selected. Darwin furthermore contends those wolves that had the best chance of survival would likewise be more likely to leave offspring and that some of their young would inherit some of their traits and, after much time and many repeated instances, there could emerge a new variety of wolf.

What function does this and the other thought experiment *illustrations* that appear in Chapters 6 through 9 play within the overall structure of the *Origin*? In a 1991 paper, James Lennox argues that these thought experiments play the role of showing the reader the explanatory potential of the theory of natural selection to explain phenomena like the increase of a variety of wolf as well as its ability to handle potentially difficult phenomena. The thought experiment with the wolf thus seems designed to help the reader see—in the

sense of imagining the thought experiment play out in their mind—the ability of natural selection to account for the selection of one trait over another in such cases. Rather than simply making the *Origin* more rhetorically effective, these thought experiments seem to show how it is *possible* that the struggle for existence could result in variety through natural selection.

Two Philosophical Thought Experiments

Thomas Hobbes and the State of Nature

Thomas Hobbes described the *natural condition* of humans in chapter XIII of his well-known 1651 work titled the *Leviathan*. This is one of the best known features of his philosophy and is commonly called the *Hobbesian state of nature*. The imagined *natural condition* of humans is a thought experiment that Hobbes used to argue for an account of human nature that would serve as the foundation of his overall argument that the best arrangement for government is a commonwealth with a sovereign. In other works, Hobbes announced that his intention was to create a *science* of politics, which he claimed no one else had done before. This science of politics was supposed to offer certain conclusions like those found in the demonstrations of geometry, and the method Hobbes followed to do so required beginning with a thought experiment. In *De corpore* Chapter VII, Hobbes provided another thought experiment at the foundation of his first philosophy, geometry, and natural philosophy, which thought experiment is similar to the *natural condition* thought experiment. This entry, however, will examine only the thought experiment that he provides in his science of politics.

One of Hobbes's primary aims in *Leviathan* was to show that the only way to achieve lasting peace was to form a commonwealth governed by a sovereign with absolute, or nearly absolute, authority that would range over all aspects of public life, including religion. To show that this end would provide peace, in Chapter XIII of *Leviathan* Hobbes asked the reader to imagine a situation in which civil relationships do not exist at all and where there is a relative scarcity of resources. This imagined situation, Hobbes admitted, never existed and never has existed in such a

pure form, although he posited that there are some places where the conditions are close to it, and he also suggested that there is such a state internationally between sovereigns. Even with such limited actualization of this state, Hobbes intended this imagined scenario to teach the reader about human nature.

Considering humans in such a state exposes, Hobbes suggested, that there is a fundamental equality shared by all humans. This equality, however, is not what one might initially expect: It is a bare equality consisting only in the ability of any human to kill any other human, whether through strength or cunning. The next feature of humans in such a state is that owing to their equality they possess hope that they can achieve whatever it is that they desire, which desires are understood in terms of appetite and aversion. Next, Hobbes suggests that these desires would become frustrated when multiple parties desire the same thing. This leads to what Hobbes calls *diffidence* and a situation in which each person takes for themselves whatever is in their power to take. This is a horrible state to be in—as the reader undergoing the thought experiment can readily surmise—and those in it are filled with anxiety and grief because each person is always worried that an attack from another may come at any time. Such a state also lacks certain other features, which Hobbes holds shows us what is *natural* and what is artificial. For example, from considering this state, Hobbes argues that there is no justice or injustice present in this natural state using the following claims: In humans' natural state, there is no common power over all humans; without a common power, there is no law; and where there is no law, there is no injustice.

The thought experiment concludes at the end of *Leviathan* 13 by identifying three final features of all humans in the imagined state of nature. Each of these humans is moved by three passions: the fear of death, the desire for the things necessary for commodious living, and the hope that those necessary things can be acquired. Hobbes thus intended the thought experiment to demonstrate the dire situation of humans without civil relationships alongside each human's natural desire to leave that state because of these three passions. Hobbes meant the reader to agree with these claims because of their inherent plausibility,

made more vivid in the form of a thought experiment. The thought experiment thus shows the imaginer *what* their own life would be like without a sovereign, and the chapters following chapter XIII aim at showing *how* to escape that horrible state.

John Locke's Account of Personal Identity, and a Criticism From Thomas Reid

John Locke also offers a thought experiment related to the natural state of humans, but this section of the entry will focus on another of his thought experiments that has generated controversy. Indeed, Locke's account of personal identity was the subject of scholarly debate by his contemporaries, but it is also relevant to contemporary philosophical debates related to consciousness, memory, identity, and the notion of the self. To motivate his understanding of personal identity, Locke offers a brief thought experiment that considers the possibility of a human's *spirit* or *soul*,—which is, in some ways, in contemporary terms equivalent to their conscious mind—leaving its body and inhabiting another one. Such a thought experiment may have seemed like pure science fiction or mere religious speculation in the past, but recent advances in artificial intelligence may make the prospects of such consciousness transfer seem more plausible than it once did.

This thought experiment, which Locke offers in *An Essay Concerning Human Understanding* (II XXVII.15), begins by asking the reader to imagine that the soul of a prince “carrying with the consciousness of the Prince's past Life” enters the body of a shoe cobbler after that individual's spirit has left it. Were this to happen, Locke argues that the prince-in-cobbler's-body would be responsible for all and only the prince's past actions. Whatever good or ill that individual had done during his lifetime would be the responsibility of the prince alone. Locke asserts, though, that no one would consider this new prince-in-cobbler's-body to be the same *man* since “the Body too goes to the making of the Man.” Locke recognizes that, of course, to all others around him the prince-in-cobbler's-body would appear to be just the same cobbler as before. Nevertheless, the prince's soul in its new body would recall all the Prince's past actions and would be responsible for them.

For Locke, this thought experiment, wherein the reader imagines the transfer of a human's conscious soul from one body to another, was partly meant to show how it was possible for a *person* (Locke's term of art in this context) to be judged by God for their past actions even though they would not appear for such judgment in exactly the same body as the one in which they did those acts. In other words, Locke was concerned to make it intelligible that persons could be judged by God even though at their bodily resurrection they would not inhabit the exact same body in which they committed those acts. This link to judgment is why Locke later identifies *person* as a *forensic term* that applies to actions and their merit, and he states that it "belongs only to intelligent Agents capable of a Law, and Happiness and Misery" (II XXVII.26). Thus, according to Locke, and supported by this thought experiment, consciousness and memory are necessary for maintaining personal identity over time.

In Locke's own time, and to the present time, there were many criticisms offered of Locke's reliance upon consciousness and memory as criteria for understanding the nature of a person. What if one's memory fails? Does personal identity, and thus responsibility, also disappear? Locke himself recognized that there were potential difficulties for this reliance upon memory and considered, for example, how the law considers the drunk and the sober individual to be the same person (and issues punishments to sober person for actions that they committed while drunk), while such is not the case according to his own account. Locke responded to such worries by admitting that human laws are limited in some respects and noted that his interest in personal identity and responsibility related to the *great Day* of judgment wherein individuals would not be held responsible for what they did not remember (II.XXVII.22). Beyond Locke's own interests in understanding how judgment at the resurrection could be made intelligible, modern readers can think about whether memory and consciousness are the appropriate criteria for personal identity, and if so, whether they should be basis for attributions of moral praise and blame generally. Do blame or praise attach only to acts committed by the individuals who remember doing them?

A well-known objection by Thomas Reid criticizes Locke's account by offering its own thought experiment, designed to pinpoint difficulties with thinking about personal identity as necessarily linked with consciousness and memory. Reid asks the reader to consider a *brave officer* at three moments in his life: when he was punished as a schoolboy for stealing from an orchard (Moment 1), when he took the standard from an enemy during a battle (Moment 2), and when he was made a general later in life (Moment 3). According to Reid's objection, the reader is asked to imagine that at Moment 2, the individual remembers Moment 1, but that at Moment 3, the individual remembers Moment 2 but forgets Moment 1. As Reid puts it, such an individual at Moment 3 would have "absolutely lost consciousness of his flogging" (Moment 1). Using this thought experiment, Reid draws the following consequence: since Locke's understanding of personal identity requires consciousness and memory for identity to be maintained, the general both is and is not the same person who was punished at school for stealing from the orchard. The worry is that that Locke's account of personal identity results in a contradiction.

Further Reflection on the Nature of Thought Experiments

How would the two primary camps, already mentioned, understand these four prominent examples of thought experiments? Arguing on the side of the *a priori* camp, in Chapter 5 of his 2011/1991 book James Robert Brown calls the Galileo example a "Platonic Thought Experiment" because it both destroys the Aristotelian view and develops its own new account. Brown argues that it transcends experience because it does not involve any new empirical data being added, because the new theory is not deduced from existing data, and because the new theory involves more than a simple adjustment to the old theory. The empiricist camp, however, would criticize such an attempt, as Norton does in a 2004 chapter, insofar as it relies upon the notion that there are Platonic Laws. How could one justify the claim that it is possible to perceive such a realm of Laws and, even if one could, how could one know that that perception is accurate?

The placement of Darwin's thought experiments within the *Origin* seems to weigh in favor of the empiricist camp's understanding of them. Darwin's evidence is provided through varied observations from his own work as well as that of others, and the aim of his *illustrations* is to help readers imagine the range of ways selection could act, thus showing the explanatory potential of the theory. In other words, empirical data form the evidentiary core of the argument, and the illustrations build upon readers' existing tacit knowledge of the external world. Likewise, the Hobbes and Locke examples both seem to be more intelligible by thinking of them through the empiricist camp's understanding of thought experiments. If we view each figure through their own epistemology, which posits that there are no ideas apart from what comes through sense experience, whether external sense or internal reflection, both Hobbes and Locke would see the role of these thought experiments as making explicit what humans already know from experience. However, it is worth noting that such knowledge gained for each figure is in some sense *beyond* their everyday experiences. As mentioned already, Hobbes admits that there has never been any time like the natural state of humans, and Locke would not have seriously entertained the claim that any of his readers had witnessed or undergone the transfer of a soul from body to another. Nonetheless, both of these thought experiments draw upon experience even as they extrapolate from it.

Marcus P. Adams

See also Biology, Evolutionary; Experiment, Theory of; Laws of Nature; Mental Models; Physics, Antiquity; Physics, Renaissance; Reasoning

Further Readings

- Adams, M. P. (2019). Hobbes's laws of nature in *Leviathan* as a synthetic demonstration: Thought experiments and knowing the causes. *Philosophers' Imprint*, 19, 1–23.
- Brown, J. R. (2004). Why thought experiments transcend empiricism. In C. Hitchcock (Ed.), *Contemporary debates in philosophy of science* (pp. 23–43). Blackwell.
- Brown, J. R. (2011). *The laboratory of the mind* (2nd ed., 1st edition published in 1991 by Routledge). Routledge.
- Brown, J. R., & Yiftach, F. (2019). Thought experiments. In E. N. Zalta (Ed.), *The Stanford encyclopedia of philosophy*. Retrieved September 1, 2021, from <https://plato.stanford.edu/archives/win2019/entries/thought-experiment/>
- Darwin, C. (2003). *On the origin of species. A facsimile of the first edition*. Harvard University Press. (Original work published 1859)
- Frappier, M., Letitia, M., & James, R. B. (Eds.). (2013). *Thought experiments in philosophy, science, and the arts*. Routledge.
- Galilei, G. (1974). *Two new sciences: Including centers of gravity and force of percussion* (S. Drake, Trans, 2nd ed.). Wall & Thompson.
- Gooding, D. C. (1992). *What is experimental about thought experiments*. PSA: Proceedings of the Biennial Meeting of the Philosophy of Science Association Vol. II: Symposia and Invited Papers, pp. 280–290.
- Gordon-Roth, J. (2021). Locke on Personal Identity. *The Stanford Encyclopedia of Philosophy* (E. N. Zalta, Ed.). Retrieved September 1, 2021, from <https://plato.stanford.edu/archives/spr2020/entries/locke-personal-identity/>
- Hobbes, T. (2021). *Leviathan* (N. Malcolm, Ed.). Oxford. (Original work published 1651)
- Horowitz, T., & Gerald, M. (Eds.). (1991). *Thought experiments in science and philosophy*. Rowman & Littlefield. Out of print but freely Retrieved from September 1, 2021, <http://philsci-archive.pitt.edu/3190/>
- Ierodiakonou, K., & Sophie, R. (Eds.). (2001). *Thought experiments in methodological and historical contexts*. Brill.
- Lennox, J. (1991). Darwinian thought experiments: A function for just-so stories. In T. Horowitz & G. Massey (Eds.), *Thought experiments in science and philosophy* (pp. 233–246). Rowman & Littlefield.
- Lennox, J. (2019). Darwinism. *The Stanford encyclopedia of philosophy*. In E. N. Zalta (Ed.). Retrieved September 1, 2021, from <https://plato.stanford.edu/entries/darwinism/>
- Locke, J. (1975). *An essay concerning human understanding* (P. Nidditch, Ed.). Oxford.
- Machamer, P., & David, M. M. (2021). Galileo Galilei. In E. N. Zalta (Ed.), *The Stanford encyclopedia of philosophy*. Retrieved September 1, 2021, from <https://plato.stanford.edu/entries/galileo/>
- Norton, J. (2004). Why thought experiments do not transcend empiricism. In C. Hitchcock (Ed.),

Contemporary debates in philosophy of science (pp. 44–66). Blackwell.

Reid, T. (1983). *Inquiry and essays*. In R. Beanblossom & K. Lehrer (Eds.). Hackett Publishing Co. (Original work published 1785)

Sorensen, R. (1992). *Thought experiments*. Oxford University Press.

Stuart, M., Yiftach, F., & James, R. B. (2018). *The Routledge companion to thought experiments*. Routledge.

TRUTH

Truth is a foundational concept. It has long been, and continues to be, indispensable to several domains of thought and practice, including government, law, science, and religion. Without the concept of truth, these practices would cease to be meaningful. Yet, the very idea of truth is difficult to capture. For millennia, philosophers have sought to pinpoint an essence or core property of truth, resulting in a wide range of complex theories. To this day, we lack an agreed-upon definition or theory of truth. Some philosophers have suggested that we simply discard the idea of truth altogether since there is purportedly nothing to define. Others have suggested that, rather than defining or rejecting the concept of truth, we should seek instead to understand what difference truth makes to everyday communicative practice. The result is that truth remains as controversial a topic in philosophy as ever. The entry traces the concept of truth from its roots in classical Greece through the coherence theories of 19th-century idealism, the work of American pragmatists C. S. Peirce and William James, the minimalist and deflationary views of more recent philosophers such as Alfred Tarski, and the links between knowledge, truth, and power as seen by continental philosopher Michel Foucault.

Correspondence Theory

The oldest and the most intuitive idea of truth is the correspondence theory, first proposed by Aristotle. In his *Metaphysics*, Aristotle asserts, “To say that that which is, is not, or that which is not is, is a falsehood; and to say that that which

is, is, and that which is not, is not, is true” (Aristotle, Γ 7.27) Aristotle’s assertion captures the idea that a true claim bears some sort of relationship with reality. While simple and commonsensical at first, a closer look reveals the difficulty of pinning down certain key concepts in this idea with precision. So, for example, what exactly is the nature of this correspondence? What corresponds with what? Is it a sentence, claim, or proposition that corresponds with reality? Is it an idea, image, or representation? It is a belief or state of mind? Consider also the nature of reality. When we refer to *reality*, are we referring to facts? If so, how are we to understand what counts as a fact? The very concept of facts is more problematic than it at first seems. Are facts impersonal items that we can collect, like rocks and leaves? Or do they have a social genesis? If the former, then we need to explain what facts are and how they can arise in the first place. If the latter, then truth would presumably correspond to something other than facts since our intuitive understanding of truth is that of something independent of human subjectivity. Thus, although many philosophers have historically endorsed the correspondence theory, they have nonetheless endorsed different versions of it: metaphysical and semantic. Despite this distinction, the correspondence theory is seen today as a decidedly metaphysical view; that is, a view that makes a strong ontological claim. While some philosophers still hold to the correspondence theory, they have nonetheless become a minority, as metaphysical claims are no longer as widely accepted in philosophy as they were in the past.

Coherence Theory

In response to the correspondence theory, the British idealists of the 19th century developed the coherence theory of truth. According to this theory, truth is a function of the internal coherence of a given belief system. A variation of this idea holds that a claim is true if it is consistent with other claims within the system. So long as that system allows us to make sense of the world around us, its degree of internal coherence can be understood as its measure of truth. The significance of this idea can be illustrated with an example from

science. If, for example, a scientific theory enables us to make sense of natural phenomena, and to conduct experiments and make predictions with meaningful results, then to the extent that the theory is internally coherent, then to that extent it is true. The coherence theory thus avoids the metaphysical problem of determining the nature of the relationship between claims or mental states and reality.

Although superficially appealing, the coherence theory nonetheless violates a common intuition about truth. If truth is a function of the coherence internal to a given belief system, then truth is relative to such systems. We can easily imagine rival belief systems with comparable degrees of internal coherence. Are they both equally true? What if they lead to different conclusions? Can incompatible conclusions resulting from competing theories both be true simply on account of the internal coherence of those theories? Truth would seem to be timeless and objective, not context-bound and relative. The coherence theory tells us more about belief systems than the phenomena they purport to explain. A second problem concerns the possibility of an incorrect belief cohering perfectly with other beliefs in a given system. If introducing an incorrect belief can still allow for coherence, then the coherence would seem to be a dubious criterion of truth.

Pragmatism: Classical and Contemporary

Recognizing the difficulty of the coherence theory, classical pragmatists such as Charles Sanders Peirce and William James took the view that truth can be measured in terms of its usefulness. On this view, the value of a given theory or belief system can be measured by how much it can deliver. This approach to truth enables us to differentiate between rival, internally coherent theories by determining which theory can produce more results. The theory that gives us more results is the truer theory. The gist of the early pragmatist view of truth is captured in James's assertion that "the true' is only the expedient in our way of thinking, just as 'the right' is only the expedient in our way of behaving." Although there is much debate about what James really meant by this vague

assertion, a common interpretation is that truth is simply *what works*. The practical appeal and non-metaphysical character of the classical pragmatist view notwithstanding, early analytic philosophers like Bertrand Russell strongly rejected this view on the grounds that it reduces truth to whatever is contingent and self-serving, and fails to specify criteria for *what works*. For example, astrology and homeopathy *work* for many people. Does that mean that either of them is true? This difficulty has led later pragmatists to rethink what Peirce and James were getting at and to offer a new interpretation and more plausible view of truth.

The neo-pragmatist philosopher Richard Rorty has been the strongest advocate in the English-speaking world for discarding the concept of truth altogether. Rorty follows Donald Davidson's view that truth is indefinable because it is a basic concept that cannot be broken down into more basic terms. Rorty further contends that truth is a metaphysical concept, like *God*, and that committing ourselves to truth is akin to submitting to God, that is, to something illusory. According to Rorty, getting rid of the concept of truth would make no difference to our social practices, like government and science. We could get by just fine without it. Huw Price has argued in turn, however, that we could not communicate meaningfully without the concept of truth, as truth implies the possibility of improving our existing claims and beliefs.

The philosopher Cheryl Misak has recently built upon certain passages from Peirce's writings to offer a new pragmatist view of truth. On her Peircean approach, truth should be understood in pragmatic terms, that is, the difference it makes to inquiry and linguistic practice. We would simply apply the term *true* to a claim or belief that appears to be indefeasible in practice, a claim or belief that we treat as capable of forever standing up to the test of practical experience. While this approach raises several questions, it nonetheless offers a non-metaphysical view of truth that answers Rorty's criticisms.

Minimalism and Deflationism

Recently, many philosophers have taken a minimalist and deflationary view of truth, regarding the truth predicate as having little meaning or

purpose other than to add emphasis to an already existing assertion or as shorthand for lengthier strings of assertions. On this approach, to say “It is true that the king of France is bald” is simply to add greater emphasis to the assertion “The King of France is bald.” Alternatively, to say “Everything in Book X is true” is taken to be a shorthand for endorsing everything asserted in Book X. Minimalism and deflationism do not treat the truth predicate as having no meaning or function, but rather see it as serving a limited linguistic purpose, one that still plays an important role in inquiry and communication.

The logician Alfred Tarski offered a kind of definition schema for the concept of truth that serves as the foundation for deflationist accounts of truth. According to Tarski, we can interpret the expression *is true* as a connection between an assertion and a state of affairs, and we can remain silent on the question “what is truth?” His account goes like this: All there is to an assertion that “grass is green” is true is that grass really is green *and* that if grass really is green then “grass is green is true.” This is expressed as:

“grass is green” is true if and only if grass is green.

According to this account there is no more to the truth concept than the collection of all possible instances of the schema.

Foucault, Truth, and Power

Whereas analytic philosophy has primarily focused on the meaning and pragmatic function of truth, the continental philosopher Michel Foucault has focused on the link between knowledge, truth, and power. Drawing from Friedrich Nietzsche, Foucault takes a much more expansive and historical view of truth. Rather than analyzing pragmatic function apolitically, Foucault proposes looking at all knowledge in terms of *truth regimes*: systems

of thought that vie with each other for institutional power, dominance, and influence. On this approach, truth is an effect of the competition for power: ideas and beliefs we treat as timeless and ahistorical, having lost sight of their historical genesis and character. Foucault’s genealogical view sees all truth and knowledge in terms of radical contingency. Given his historical and cultural approach to truth, Foucault’s genealogical method has been widely adopted in the humanities and social sciences to examine a wide range of systems of thought with the intention of understanding their power implications.

Jason Hannan

See also Epistemology; Evidence; Explanation; Generalization; Intuition; Knowledge; Metaphysics; Pragmatism

Further Readings

- Blackburn, S. (2005). *Truth: A guide*. Oxford University Press.
- Blackburn, S., & Simmons, K. (Eds.). (1999). *Truth*. Oxford University Press.
- Crivelli, P. (2004). *Aristotle on truth*. Cambridge University Press.
- Davidson, D. (1996). The folly of trying to define truth. *The Journal of Philosophy*, 93(6), 263–278.
- James, W. (1911). *The meaning of truth*. Longman Green and Co.
- Misak, C. (2004). *Truth and the end of inquiry*. Oxford University Press.
- Price, H. (2003). Truth as convenient friction. *The Journal of Philosophy*, 100(4), 167–190. <http://www.jstor.com/stable/3655652>
- Rorty, R. (1995). Is truth a goal of enquiry? Davidson vs. Wright. *The Philosophical Quarterly*, 54(180), 281–300. <https://doi.org/10.2307/2219651>
- Taylor, C. (1984). Foucault on freedom and truth. *Political Theory*, 12(2), 152–183. <https://www.jstor.org/stable/191359>

U

UNDERSTANDING

The way people think has long intrigued philosophers: *metaphysically* (regarding the nature of human intelligence) and *epistemologically* (regarding the nature of human reasoning) and, more recently, *psychologically* (regarding the scientific verification of naturally occurring phenomena in human thought). Theories of understanding try to explain the changes in the individual as well as describe the categorical differences across types of understanding.

Understanding moves us beyond the knowledge of isolated propositions to a more integrated grasp of the connections, relations, and significance of what we know. Indeed, we tend to think that knowledge, while important, is only the beginning of true understanding, which requires situating that knowledge within our entire cognitive economy. People often use terms of sensory perception in relation to developing understanding, such as “Now I see!”, but it refers to the satisfactory resolution of meaning concluded by the individual. An ancient admonishment highlights the difference between exposure and comprehension: “Hear now this, O foolish people, and without understanding; which have eyes, and see not; which have ears, and hear not” (Jeremiah 5:21). Understanding will integrate *experience discovered* (via personal perceptions) and *knowledge received* (from others’ reports of their own perceptions).

Graduated Understanding

Instead of a binary difference between understanding and misunderstanding, there is a common assumption that understanding can increase gradually. This may be explained by maturation of the individual, more complexity of one’s context, or a sequence within the nature of understanding. Although the arrival of the understanding may be experienced as an “aha” moment of inspiration, it is generally regarded as the cumulative result of many factors.

A lack of understanding can thus be attributed to insufficient opportunity to experience and process the concept in a meaningful context, with the additional factors of the individual’s capacity to do so and whether there is curiosity or interest. Because each of those factors will vary, the understanding is also variable in its pace and complexity. One’s degree of understanding may be described as incomplete, partial, or biased, but never permanently fixed the way one’s capacity might be determined. Understanding is expected to change, hopefully ever toward greater wisdom but nonetheless vulnerable to deception and ignorance of emerging new information.

Understanding is seen as a continuous process, with knowledge as the product of a degree of sophistication in its development. Knowledge tends to be accepted if it has been justified as true and is believable to society in general, but understanding is used as an indicator that the individual no longer experiences confusion. People are considered knowledgeable if they are competent in

the retrieval and application of information, but the people may not be regarded as having understanding if broader consequences and multiple perspectives are not apparent. Thus, many think of understanding not only as a continuous process but as part of a continuum of thought.

The active process of understanding is inherent in the taxonomy of the cognitive domain developed by the educational psychologist Benjamin Bloom (1913–1999) and colleagues, and in its revision half a century later. The taxonomy included three domains: *psychomotor* (movement), *affective* (feeling), and *cognitive* (thinking). The cognitive taxonomy was developed as a tool for assessment; that is, a way to measure the degree of cognitive proficiency as a result of different learning techniques. There have been many models of cognition, but Bloom's Taxonomy of Cognition (named for the lead researcher on the team) has become a standard way for educators to analyze learning objectives and to categorize student performance.

As shown in Table 1, two dimensions intersect: four types of knowledge and six types of thinking, or cognition. Understanding, the second cognitive process in the table, occurs in each knowledge dimension of *facts*, *concepts*, and *procedures*. The fourth dimension of knowledge is *metacognition* or self-awareness. Bloom's Taxonomy helps teachers be more accurate in their assessment of students' learning, and for that reason, the categories of the cognitive process dimension, such as *remember* and *understand*, are considered too vague for a learning objective. Instead, more specific verbs like those listed in Table 1 are used as indicators of mastering the level of cognition.

As used here, understanding is considered more advanced than simple factual recall, and it is also

considered necessary for the more sophisticated processes of analysis, decision-making, and creativity. The defining characteristic for this level is to be able to explain the meaning and context of the remembered information. The action verbs associated with the level in Table 1 suggest that understanding is in response to novel situations and in anticipation of future events. This taxonomy was proposed as a hierarchy from low to high levels of thinking, but there is not widespread agreement that the various types of thinking must occur in this sequence.

Understanding and cognition are not synonymous. Neuropsychological abilities include various aspects of thinking and reasoning, that is, understanding exhibited by *verbal memory*, *executive function*, *inhibitory control*, *processing speed*, *working memory*, and *vocabulary*. These cognitive functions are considered domain-general because they are used for all domains of experience. Researchers have devised reliable methods to test and measure these abilities, producing a large volume of literature informing our understanding of understanding.

Psychologist Howard Gardner's (1949–) theory of *multiple intelligences* expands this to identify many ways of knowing, challenging the traditional view of intelligence as defined by traditional academic proficiency. The separate modalities are not stable traits that limit individuals' capacities to understand, but rather a framework for many ways in which each individual can access information and process it meaningfully. According to Gardner, the frames of mind manage to span all aspects of human interaction with experience: the *cognitive* (logical–mathematical and verbal–linguistic), the *psychomotor* (visual–spatial, body–kinesthetic, and

Table 1. Verbs Representing Each Cognitive Process Dimension in Each Knowledge Dimension According to Anderson and Krathwohl's *A Taxonomy for Learning, Teaching, and Assessing: A Revision of Bloom's Taxonomy of Educational Objectives*

<i>The Knowledge Dimension</i>	<i>The Cognitive Process Dimension</i>					
	<i>Remember</i>	<i>Understand</i>	<i>Apply</i>	<i>Analyze</i>	<i>Evaluate</i>	<i>Create</i>
Factual	List	Summarize	Classify	Order	Rank	Combine
Conceptual	Describe	Interpret	Experiment	Explain	Assess	Plan
Procedural	Tabulate	Predict	Calculate	Differentiate	Conclude	Compose
Metacognitive	Custom	Execute	Construct	Achieve	Action	Actualize

rhythmic–harmonic), and the *affective* (interpersonal/social and intrapersonal/metacognitive).

Gardner insisted that our understanding of intelligence is ongoing. He and his colleagues introduced an eighth one focusing on understanding systems in context (naturalist) and at least one more proposed (existential). The pedagogical implication is that *multiple modalities* should be used to both develop and demonstrate understanding. The theory is supported by recent research confirming that specific neural activity is associated with each intelligence.

Developmental Stages of Understanding

Whereas Bloom's cognitive taxonomy categorizes different types of thinking, other theories of understanding focus on the person and how patterns of understanding will change across time. Eric Erikson (1902–1994) proposed eight psychosocial stages, each of which depends upon understanding, that is, the healthy resolution of predictable crises. Unlike the innate psychosexual stages of Sigmund Freud (1856–1939), Erikson's theory suggested that lack of understanding explains dysfunction occurring later in life, thus emphasizing the skills needed to navigate the landscape of inner conflict. This aligns with John Dewey's (1859–1952) hypothesis that understanding occurs only if there are problems to solve.

The capacity to understand clearly changes, prompting many theories to explain and predict it as a natural phenomenon. Jean Piaget (1896–1980), the Swiss developmental psychologist, focused on intellectual development, coining the term *genetic epistemology* for the changes he observed in children's capacities to understand different phenomena. He proposed that we respond to *cognitive dissonance* by either adjusting the new information to accommodate our schema for understanding the world or we adjust our schema in order to assimilate the new information. His theory suggested that interaction with objects was necessary for cognitive development.

The Soviet theorist Lev Vygotsky (1896–1934) outlined a sociocultural psychology featuring the *zone of proximal development* (ZPD). The ZPD is the expanded potential for learning found in being near someone with a larger knowledge base. The interaction itself was not the object of

understanding, but rather the context for improved understanding of interest to both people. This theoretical construct can help explain the well-documented success of *cooperative learning groups*, an instructional method requiring interaction between students sharing a task of solving a problem. American psychologist Carl Rogers (1902–1987) popularized the importance of a personal relationship to be established before a client could benefit from a therapist's guidance or a student could benefit from a teacher's influence. Understanding is therefore subject to a noncognitive context. Current pedagogical practice recognizes this as *social–emotional learning* that aligns with Bloom's affective domain.

Both Piaget and Vygotsky are considered *constructivists*, meaning that their theories emphasize the individual's understanding as personally distinctive from others' understanding. Constructivism as well as progressive pedagogy, emphasizing active engagement in real-world problems, have profoundly influenced schools to become more learner-centered and therefore more *developmentally appropriate*. Teachers are expected to be familiar with norms of cognitive development in order to maximize instruction for their students' understanding. For instance, children aged 5–7 begin to grasp cause-and-effect relationships plus more logical concepts of time and space, but their source of knowledge will still be concrete experience.

Canadian educational philosopher Kieran Egan (1942–) proposed an *imaginative* theory of learning that recognized five categories of understanding: *somatic* (physical functions and emotions), *mythic* (binary opposites), *romantic* (transcending one's own limits), *philosophic* (patterns and principles), and *ironic* (capacity for paradox and alternative thought). Each of these *understandings* is regarded, like Gardner's Multiple Intelligences, as different ways to arrive at truth and meaning.

Personal Understanding

As the preceding sections suggest, understanding is a personal event, akin to making sense of the world. Viktor Frankl (1905–1997), psychoanalyst and Holocaust survivor, identified humans' *search for meaning* as the primary function of life, explaining individuals who failed to thrive in the face of changed circumstances. Their personal meaning had

changed. Similarly, William G. Perry (1913–1998) did not assume individuals' stable character traits would predict ethical behavior. Instead, he focused on their progress through four stages of intellectual and ethical development, namely *dualism*, *multiplicity*, *relativism*, and *commitment*, as individuals progress through a gradually more complicated life. His scheme has been the basis for most subsequent ethical theories with multiple dimensions. A different view was presented by Mary Field Belenky (1933–2020) and colleagues, who challenged Perry's scheme first of all because it was based on studies limited to men. *Women's Ways of Knowing* added a novel first stage of *silence*, that is, being disconnected from knowledge. It reorganized the remaining stages of *received*, *subjective*, *procedural*, and *constructed* knowledge, the last of which would integrate all voices. Perry's scheme focused on the nature of knowing in response to changing contexts, while Belenky, in addition to women's perspectives and a broader group under study, focused on the sources of knowing.

Naomi Jeffery Petersen

See also Epistemology; Explanation; Inference; Intuition; Knowledge; Reasoning

Further Readings

Anderson, L. W., & Krathwohl, D. R. (2001).

A taxonomy for learning, teaching, and assessing:

A revision of Bloom's Taxonomy of Educational Objectives. Longmans.

Bernstein, D. M., Thornton, W. L., & Sommerville, J. A. (2011). Theory of mind through the ages: Older and middle-aged adults exhibit more errors than do younger adults on a continuous false belief task.

Experimental Aging Research, 37(5), 481–502.

Bransford, J., Brown, A., & Cocking, R. (Eds.). (2004). *How people learn: Brain, mind, experience, and school.* National Academy Press.

Egan, K. (1997). *The educated mind: How cognitive tools shape our understanding.* University of Chicago Press.

Gopnik, A. (2010). How babies think. *Scientific American*, 303(1), 76–81.

Lipton, P. (2009). Understanding without explanation. In H. de Regt, S. Leonelli, & K. Eigner (Eds.). *Scientific understanding: Philosophical perspectives* (pp. 43–63). University of Pittsburgh Press.

Nichols, S. (2000). Folk concepts and intuitions: From philosophy to cognitive science. *Trends in Cognitive Science*, 8(11), 514–518.

Piaget, J., Inhelder, B., & Szeminska, A. (1960). *The child's conception of geometry.* W.W. Norton & Co.

Shearer, C. B., & Karanian, J. M. (2017). The neuroscience of intelligence: Empirical support for the theory of multiple intelligences? *Trends in Neuroscience and Education*, 6, 211–223.

Taylor, S. J., Barker, L. A., Heavey, L., & McHale, S. (2013). The typical developmental trajectory of social and executive functions in late adolescence and early adulthood. *Developmental Psychology*, 49(7), 1253–1265.



VALUES IN SCIENCE

Values in science may refer to (1) the orienting aims or constitutive values of science as such or (2) the role of extrascientific values in the development of scientific activity. A related view that might be termed “values *and* science” concerns (3) the manner in which scientific knowledge and the esteem in which it is held in a given society itself shapes values.

The first view concerns the organizing principles as well as epistemic bases of science (i.e., what counts as scientific inquiry, evidence, or knowledge). The second view concerns both (a) the role played by extrascientific values in guiding the direction of science (e.g., focusing scientific resources on determining the nature of culture, on the causes of cancer, or on extraterrestrial matter) and in conferring authority on science, which has implications for science’s prominence and centrality in a given society, and (b) the role of values in conditioning scientists’ views of the world. The latter concern has been the subject of controversy in the philosophy of science as well as in the sociology of science, and it is at the heart of the debate between positivists, or naturalists, who see science as reflecting concrete and certain reality and being more or less impervious to influence by values, and social constructivists, or relativists, who claim that science serves to concretize values held by the culture in which it is developed.

Values and the Constitutive Values of Science

Values are the evaluative dimension of the basic constellations of shared symbolic orientation that structure action and human social experience. Values underlie rules, norms, and ethical practices, and they shape patterned activities. The attribute of shared orientation makes manifest the control that values exert on human behavior.

Science, like all other regularly patterned symbolically constituted activities, is a product of a peculiar set of values or shared symbolic orientations. Science may best be understood as a *patterned activity* characterized by a particular aim or goal: to attain objective knowledge of empirical phenomena. This aim constitutes science’s central organizing principle and thus shapes the activity of science as well as the standards according to which scientific knowledge is created. This aim, moreover, indicates the set of phenomena relevant for science as well as criteria for an appropriate method to be employed in pursuit of it. The various sciences (physics, chemistry, biology, sociology, etc.) are organized around sets of phenomena that are deemed relevant for the problems that motivate their scientific inquiry and that are indicative of their respective goals. In this way, each branch of science is special. The sets of phenomena that define them reflect the value placed on what is considered as worthy of being known. The constitutive values of science are intimately related to its aim and follow from it. The criterion of objective knowledge in science’s constitutive set

of aims demands that science be value-free or value-neutral in the sense of being resistant to extrascientific values such as those that are contained in ideologies or political imperatives.

The values of science denote those orientations that individual scientists and the scientific community should have if they are practicing good science. The reality of scientific practice, however, is indeed much more complicated. The sociology of science seeks to understand not only the value orientations of the patterned activity of science but also how practicing scientists interpret, transform, or stray from these values in practice. The line between good science and bad science is drawn by appealing to science's organizing principles, which are contained in its aim and the values that stem from it.

The core value or internal presupposition of science is the importance of solving a problem, which is to say that that which is yielded by inquiry is regarded as worth being known. The solution, moreover, must refer to empirical phenomena. Robert Merton, the esteemed American sociologist of science, argues that the term *science* refers to a number of interrelated albeit distinct concepts. One, it denotes a set of methods used to certify knowledge; two, it refers to the accumulated knowledge deriving from the application of such methods; three, it refers to a set of values and mores governing scientific activities; and, four, any combination of these. According to Merton, the set of values governing scientific activities makes manifest the following imperatives: universalism, communism, disinterestedness, and organized skepticism. *Universalism* denotes the value that the worth of scientific discovery is contained in the work itself and not in the personality of the work's creator. Criticism of science should not be ad hominem; positive appreciation of science should not be based on charismatic qualities of the scientist either. *Communism* denotes the value that work in science is communal: Great science is done, as it were, by standing on the shoulders of giants. Additionally, the value of communism demands that work be shared with the public. Privately held research betrays the values of scientific communality. *Disinterestedness* is a value that guides science toward something like sincerity in research. The aim of science is tied to the value of

the problem and solving it: This is its guiding light. Science is essentially not a means to fame, money, or power. *Organized skepticism* is the value orientation toward continual questioning and systematic testing of facts and theories. The avoidance of dogmatism stems from this value. These four features may be said to be four characteristic values of science. On this view, science manifests organizing principles that delineate it essentially from other kinds of patterned activities (such as figurative painting, imaginative literature, or religion). Honesty and objectivity are obvious values tied to those that Merton notes.

According to Karl Popper, the well-known 20th-century philosopher of science, the activity of science is characterized by the fact that its problems are formulated in such a way that demands that claims or theories be capable of refutation and that those claims are indexed to sets of empirical phenomena that are observable. A fit must be established between proposed solutions or theories and the relevant empirical facts as well as between descriptive claims about the relevant empirical facts and the objects described. It is clear that human judgment enters science during this process, not to mention at the initial stage where questions get formulated. The question is "How do scientists make judgments that are not informed by values?" The answer that has become the mainstream view over the past 50 years or so in the philosophy of science, in the sociology of science, and in science and technology studies is "they do not"—values are everywhere in science. On this view, scientists go out into the world looking for something, and out of the vast chaos of reality scientists impute an order and select questions considered to be important, evidence considered to be relevant, and tests and methods considered significant or appropriate. Popper is helpful here. He argues that an *expectation* or a *frame of reference* always precedes an *observation*. Values hence guide discovery. Popper refers to expectation-driven observation as the *searchlight theory* of science, which stands opposed to what he calls the *bucket theory*, which is something like a pure inductivist view of phenomena. No scientist can take everything into account.

With this in mind, it is clear that one of the basic value assumptions of science is that the

world is ordered. This shared symbolic orientation is a frame of reference in the Popperian sense. As the famous exchange between Albert Einstein and Niels Bohr over the development of quantum mechanics in physics made plain, one of Einstein's basic values was the belief in fundamental order: On several occasions, he expressed the sentiment that "God does not play dice with the universe." Einstein's frame of reference—his searchlight—is a value.

The Place of Extrascientific Values in Science

Values and the Direction of Science

In modern society, science is accorded positive social value. In this respect, science itself is a value. Science is not only revered but also constitutes the most authoritative form of knowledge of the current era. This was not always the case, however, nor is it a universal position: The honored position accorded science once belonged to theology and philosophy, and skepticism of science endures. The social value of science is an external condition of the activity of science itself—the social value of science neither follows from the internal logic connected with its aim or goal nor does it follow from the norms or characteristic practices of the activity itself. The positive social value accorded science is thus analytically separable from science. The positive social value of science is in this way an extrascientific factor.

Other extrascientific values bear on science, but, as has been argued by Robert Merton, Joseph Ben-David, and Liah Greenfeld, they do not affect the essential nature of the patterned activity or the criteria of evaluation internal to it. For one, there is the logical contradiction: If science were to be guided by a set of values that privileged the transcendental realm over the empirical/phenomenal realm, for instance, it would be clear that the patterned activity was not science at all but something approaching religion. As the pioneering sociologist Max Weber elucidated in his methodological writings on science and in his well-known *vocation* lectures, scientific values and political values are incommensurable. Each set of values might inform the other, but one cannot be reduced

to the other. Just as the game of soccer and gridiron football are guided by organizing principles that define them, so too are science and politics. If one introduces the rules of football into a game of soccer, one is no longer playing soccer. From a sociological standpoint, when the internal logic of a patterned activity is altered, it constitutes a new phenomenon. Weber, as much as anyone has done, recognized that values cross cut all human activities—science included. The key is to maintain a value orientation in science that actively seeks to keep extrascientific values out.

The direction of science is affected by recognized values in one significant respect—that concerning the problems considered worthy of being pursued. The value accorded solving problems related to various cancers, for instance, determines which scientific research programs receive funding, the degree of social approbation conferred on the research carried out, and which lines of inquiry are to be accorded priority. The urgency of the AIDS epidemic influenced the direction of research in biology, taking biology in a direction different from that of its previous course, and political developments (consider World War II and the Cold War) shaped the direction of research in physics.

The degree of development and prestige of science is under the control of values. As Merton argued in his classic text *Science, Technology, and Society in 17th Century England*, among the conditions for the explosive growth of science was the set of religious values of science's early proponents in England. Science was seen as a means of discovering God's intentions and the constitutive qualities of the natural order. More than this, the growth of scientific knowledge was envisioned as demonstrating the glory of God. Isaac Newton himself was a scientist who was motivated to pursue science in part by his religious values. Liah Greenfeld has complemented and amended Merton's thesis by demonstrating that the birth of nationalism in England helped bring about a shift in values regarding science, making the pursuit of science a national priority and an object of prestige. The pursuit of national prestige—a status value—constitutes in Greenfeld's thesis a significant part of the explanation concerning the sustained development of science. The pursuit of prestige is also a contributing

factor at the level of individual scientific achievement. The rush to publish one's results first, the expectation of receiving due credit for a discovery, and the hope of receiving a prize for one's work are all motive forces of science deriving from extrascientific values.

Values and Scientific Worldview

The role played by values in science has been and continues to be the subject of controversy. This controversy may be understood as a debate over two polar positions: one, that science is by definition value-free and only concerns the so-called facts of the world; and, two, that science is just a particular instantiation of human culture in general, and the view it offers on reality is no less conditioned and shaped by the complex of symbolic frameworks and shared symbolic orientations underlying values through which its activities and resultant claims are realized. These polar positions are positivism (or naturalism) and constructivism (or social constructivism/relativism). In its extreme version, positivism sees science as having privileged access to the facts of reality plain and simple, and general laws hence are drawn from facts. For its part, social constructivism sees science as one among many types of meaning-making or symbolic activities; in this way, the arbitrary standards typical of meaning-making processes are seen to impugn the credibility of the claims to value-freedom in science.

The view that conceives of science as dealing basically with observable and measurable facts and describing putative universal laws or empirical regularities without values or other extrascientific factors coming into play is termed *positivism*. Positivists adhere to a nomothetic ideal of science, which is to say an ideal that conceives of the world as under the control of general laws. Like all movements of thought, positivism encompasses a range of viewpoints. A uniting characteristic of positivist views is a strong belief that scientific facts are plain, unadorned, universally intelligible, and so forth. Based on facts thus understood, scientists draw out laws by generalization and extrapolation. Positivism not only rejects metaphysical explanation but also the view that social values play a determinative role in science. This view was the dominant scientific perspective

throughout much of the 19th and 20th centuries. Positivism was never the mainstream view in the sociology of science. It is no longer the mainstream viewpoint in the philosophy of science or other disciplines that study science and the production of scientific knowledge, although variants of this view continue to be held among the masses as well as among many practicing scientists and some philosophers.

The view that stands in contradistinction to positivism is *social constructivism*. A turning point in the commonplace positivist view of the relation between values and science came about with the publication of Thomas Kuhn's book *The Structure of Scientific Revolutions*. Although many philosophers and other scholars had advanced views that were destabilizing to the positivist foundation, Kuhn's book had a wide appeal that served, as it were, to tip the scale to the side of social constructivism. Kuhn argued that—contrary to the conventional wisdom that the logic of scientific development proceeded according to an internal momentum of the state of theory and that scientific knowledge was incremental, cumulative, and hence progressive—scientific theory was actually characterized by periods of systematic development (he called this normal science) that are interrupted by so-called revolutionary changes in its orienting paradigms. A *paradigm* is an exemplary finding or theory in which something like a strong consensus or shared positive appreciation obtains, and one might regard a paradigm as functioning axiomatically in times of normal science. Examples of paradigms offered by Kuhn include Aristotelian motion or the mathematical view of the electromagnetic field offered by James Maxwell. A major implication of Kuhn's view is that paradigms are accorded value and function as standards by which the adequacy of a theory or finding is judged. This view stands in stark opposition to the view of scientific development in which evidence is as if objectively measured against the theory according to methodological rules. On Kuhn's view, the structure of theoretical logic as represented by paradigms is regularly overturned and replaced by new theories that, since they follow different paradigms, are incommensurable.

The social constructivist view of science, which argues that values influence science implicitly

(e.g., in the construction of facts or in the determination of significant phenomena) and that broader social values shape the outcomes of scientific debate, interprets Kuhn as offering support to their view. Social constructivism suggests a view considerably stronger than the view that extrascientific considerations merely direct science in one direction or another: It argues that the facts and theories of science themselves are under the control of extrascientific values. The perspective of social constructivism is developed in the philosophy of science, sociology of science, science and technology studies, the sociology of scientific knowledge, as well as other disciplines and movements (e.g., feminist theory). A central contention of the social constructivist view is that values matter for scientific outcomes—change the values, and the outcomes change.

Values and Science

Scientific discoveries have played a role of their own in shaping the values of the wider culture. One important role in this regard concerns science's privileged status as a form of knowledge. Scientific explanation is regarded in modern societies as not only reliable but authoritative. Science has become the arbiter of so-called truth. As demonstrated by the public debates concerning the value of Charles Darwin's theory of evolution vis-à-vis the Christian biblical view on the origins of human life as well as the ontological questions these debates have given rise to, science is a force for shaping values outside its own sphere of knowledge production. Scientific claims have affected the values concerning standards of evidence in courts of law, the direction and course of medical care, and the allocation of funding to nations in need. What might be the most salient contemporary case demonstrating science's effect on social values concerns the claims regarding climate change made by the scientific community and the evaluation of values it has given rise to. The case of climate change also draws attention to the role played by scientific technology in shaping values by way of the ethical debates new technologies have prompted. The advent of the birth control pill, for instance, has played a role in shaping values regarding family planning, personality responsibility, and the advancement of women.

New medical technologies such as vaccination and organ transplants have led to evaluations of values so as to determine thresholds of acceptable risks. What complicates cases such as these is that determinations of acceptable risk often draw from science for evidence. Such examples illustrate that values crosscut human activities, yet the values characteristic of various activities remain distinct.

Eric Malczewski

See also Fact Versus Theory; Philosophy of Science; Scientific Realism

Further Readings

- Ben-David, J. (1971). *The scientist's role in society: A comparative study*. Prentice-Hall.
- Greenfeld, L. (1987). Science and national greatness in 17th century England. *Minerva*, 25(1–2), 107–122.
- Greenfeld, L. (2006). Science and literature as social institutions. In Greenfield, L., *Nationalism and the mind: Essays on modern culture* (pp. 145–161). Oneworld Publications.
- Malczewski, E. (2013). This is social science: A “patterned activity” oriented to attaining objective knowledge of human society. *The Journal of Classical Sociology*. <https://doi.org/10.1177/1468795X13495124>
- Malczewski, E. (2016). The value(s) of nature. *American Journal of Cultural Sociology*, 4(2), 228–239. <https://doi.org/10.1057/ajcs.2016.3>
- Malczewski, E. (2019). The Weberian presuppositional analytic. *Sociological Theory*, 37(4), 363–380. <https://doi.org/10.1177%2F0735275119888258>
- McMullin, E. (1982). Values in science. *PSA: Proceedings of the Biennial Meeting of the Philosophy of Science Association*, 2, 3–28.
- Merton, R. K. (1942). Science and technology in a democratic order. *Journal of Legal and Political Sociology*, 1, 115–126.
- Merton, R. K. (1970). *Science, technology and society in seventeenth century England*. Fertig. (Original work published 1938)
- Merton, R. K. (1973a). The normative structure of science. In N. Storer (Ed.), *The sociology of science: Theoretical and empirical investigations* (pp. 267–280). University of Chicago Press.
- Merton, R. K. (1973b). Priorities in scientific discovery. In N. Storer (Ed.), *The sociology of science: Theoretical and empirical investigations*

- (pp. 286–324). University of Chicago Press. (Original work published 1957)
- Merton, R. K. (1973c). Science and the social order. In N. Storer (Ed.), *The sociology of science: Theoretical and empirical investigations* (pp. 254–266). University of Chicago Press. (Original work published 1938)
- Merton, R. K. (1996). The ethos of science. In P. Szotompa (Ed.), *On social structure and science* (pp. 267–276). The University of Chicago Press.
- Popper, K. (1972). *Objective knowledge: An evolutionary approach*. Oxford University Press.
- Popper, K. (2002). *The logic of scientific discovery*. Routledge.
- Popper, K. (2008). Science: Conjectures and refutations. In K. Popper, *Conjectures and refutations* (pp. 43–86). Routledge.
- Weber, M. (1958a). Politics as a vocation. In H. Gerth & C. Wright Mills (Eds. & Trans.), *From Max Weber: Essays in sociology* (pp. 77–128). Oxford University Press. (Original work published 1921)
- Weber, M. (1958b). Science as a vocation. In H. Gerth & C. Wright Mills (Eds. & Trans.), *From Max Weber: Essays in sociology* (pp. 129–156). Oxford University Press. (Original work published 1917)
- Weber, M. (2011a). Critical studies in the logic of the cultural sciences. In E. A. Shils & H. A. Finch (Eds. & Trans.), *Methodology of social sciences* (pp. 113–188). The Free Press. (Original work published 1906)
- Weber, M. (2011b). “Objectivity” in social science and social policy. In E. A. Shils & H. A. Finch (Eds. & Trans.), *Methodology of social sciences* (pp. 49–112). The Free Press. (Original work published 1904)

VITALISM

Vitalism is the doctrine that living organisms are essentially distinct from nonliving objects because they possess either a nonphysical property or a distinctive organization not found among nonliving objects. Historically, the vital property has been identified with a range of things, including a fluid, a spirit, or a life-giving substance. The ancient Greek physician Galen (129–216 C.E.) located the vital force in the air (*pneuma*) that is absorbed through the lungs and infused with blood in the heart. The French anatomist Xavier

Bichat (1771–1802) located the vital properties of living tissues in the form of *sensibility* and *contractility*, fundamentally irreducible properties (on par with gravity and elasticity in physics) that did not decompose. The chemist Justus Liebig (1803–1873) identified the vital force as a regulative function that controls absorption and waste release processes in plants and animals. The biologist Hans Driesch (1867–1941) identified a living force, or *entelechy*, in developmental processes. The philosopher Henri Bergson (1859–1941) argued that there must be a vital impulse or *élan vital* that is common to all forms of life and is responsible for the creation of all species.

Contemporary scientists are likely to dismiss the doctrine of vitalism for its mysterious nonmaterialistic ontological commitments. Philosopher David Hull (1935–2010) claimed that scientists and philosophers take ontological reductionism for granted, vitalism is dead, and that organisms are nothing more than atoms. Francis Crick (1916–2004), reflecting on his contribution to the helical structure of the DNA molecule to the investigation of life, is even more forceful, claiming that in light of molecular evidence, vitalists are *cranks*.

Yet, there are good reasons for taking vitalists and their motivating arguments seriously. Many vitalistic proposals (since the 19th century) were offered as a reaction to inadequacies or shortcomings of physicalistic explanations including its method of analysis, reductionism, interpretations of cause and effect, and adequacy conditions for empirical confirmation of unobservable or theoretical entities, especially for the kinds of complex systems associated with living organisms.

For instance, in reaction to Crick’s missive, the biologist C. H. Waddington (1905–1975) pointed out that physicists often postulate the existence of unobservables in order to account for peculiar observational effects or incomplete ordering schemes. A quark is considered a fundamental constituent of matter even if (at the time of postulation) it was not observed. Waddington’s point is that postulating unobservables is common practice in physics and hence should be acceptable in biology as well. In the philosophy of science, the question about how we should think about theoretical, unobservable entities is the basis of much discussion.

Another reason not to be overly dismissive of the vitalist doctrine is that not all vitalists appeal to nonmaterial substances as the basis for the distinction between living and nonliving organisms. Contemporary historians of biology have recently recognized another branch of vitalists—*enlightened vitalists*—in the 19th century who emphasized the type of articulation of the parts of organisms and their unique arrangements in the body as the locus for biology pertaining to understanding organisms. The French physician Ménéuret de Chambaud (1739–1815) suggested that the life in a body occurs as a connection of actions or an aggregate effect of interactions among various parts. He illustrated his point by comparing organismal arrangements with bee swarms, which strive together to hang on to a tree branch. Hanging on to a tree branch is not a characteristic of any individual bee, but it is what they do collectively. Likewise, life is not found in any part of a body, but it is a property that emerges from the arrangement contribution of all parts.

In sum, vitalism can be seen both as an attempt to explain the distinctness of living organisms and as a reaction to reductionistic explanations which refer to physical mechanisms and chemical structures. Vitalism emerged when some natural philosophers started to propose several distinct hypotheses that could be unified around the idea that life had inherent *vital properties* that differ from the sort of properties inorganic objects display.

Hans Driesch's Neovitalism

Hans Driesch was the most influential early 20th-century biologist to defend vitalism. In 1891, he performed experiments on sea urchins that appeared to contradict Wilhelm Roux's 1888 experiments on frog development. Driesch's initial interpretations of the experimental results were foundational for his vitalistic philosophical doctrine because Driesch did more than philosophize about the life-giving properties of organisms, he subjected his views to empirical testing.

Roux's experiment involved destroying half of the cells in frog embryos after the first stage of cellular division. Roux reported that the remaining cells continued to develop into half-embryos and conjectured that the function of each cell

played a unique part in the entire developmental process. Roux's work initiated the field of *Entwicklungsmechanik* or developmental mechanics. There is no need to postulate any vitalistic forces to explain embryonic development.

Driesch's sea urchin experiments, however, provided contradictory results: The nondestroyed embryonic cells continued to develop into full (but small) sea urchins. Driesch showed that it was possible to shuffle the blastomeres (cells produced by the first cleavage of the zygote, or fertilized egg) of early embryos with no effect upon later development. This suggested to Driesch that development does not involve a mechanistic predetermined process because any cell—regardless of where it is located in the system—had the potential to develop into a complete organism. Driesch did not deny that development involved chemical and physical processes, but he claimed that a complete explanation required the postulation of special conditions. Embryo development is not—as the practitioners of *Entwicklungsmechanik* believed—merely a matter of division of a germ that controls cellular division. Rather, development involves the response of a living organism and its immediate surrounding conditions.

On the basis of his experiments, Driesch proposed that behind living organisms there was a living force, called an *entelechy*, which would explain the seemingly anti-mechanistic behavior in developing embryos. The entelechy is present not just in embryos but in all organisms. It is a regulating agent that determines which of the various potential developmental outcomes the organism will realize. It also acts as a buffer against potential deleterious effects that organisms encounter from environmental conditions. Although the entelechy is nontemporal, nonspatial, it does not contradict any physical laws. For this reason, Driesch is often labeled as a *neovitalist*.

Vitalism and Organicism in the 20th Century

In the early 20th century, discussions about vitalism centered on two broad questions. The first is a question about metaphysics: What is the fundamental trait of living beings that allows us to draw the line between them and the inanimate world?

The second question concerns scientific methodology question in the investigation of life: If biology is a distinct science (other than physics or chemistry), what makes it distinct?

Among philosophers, Henri Bergson is among the most important vitalists around the end of the 19th century and the beginning of the 20th century. Bergson argued that there must be a vital impulse, an *élan vital* that is common to all forms of life and is responsible for the creation of all species. As a consequence, both mechanistic and traditional teleological approaches were, in his view, unable to account for changing and creativity, two characteristics necessary for evolution; Bergson advocated for a distinct form of finalism, compatible with the existence of the postulated vital impulse.

In 1914, the philosopher James Johnstone stated that organisms, when considered as a whole, provide disproof of the mechanistic explanatory paradigm. He argued that it is possible to study organisms from a purely physicochemical point of view and that reductionism to organisms to physicochemical phenomena is possible for methodological reasons (e.g., when scientists isolate parts of the organisms), but such approaches leave out something, since organisms should be considered as a whole. What is left out is a form of an *elemental agency* that would be responsible for the *direction and coordination of energies*.

For the Austrian biologist Karl Ludwig von Bertalanffy (1901–1972), an important lesson learned from developmental biology was its anti-reductionistic character. Although organic processes never run contrary to laws of physics and chemistry, life is more than a heap of physical and chemical processes. It is a particular arrangement and organization of physical and chemical processes.

By maintaining a commitment to physical processes while at the same time emphasizing a special characteristic of organic life, Johnstone and Bertalanffy were harkening back to the Enlightenment vitalism of the 19th century. Subsequent proponents of this approach are called *organicists*. Organicists reject the vitalistic doctrine that postulates the existence of nonmaterial life forces. Yet, they retain the idea that there are systematic features of organisms that distinguish them from

nonliving entities. Three defining themes of the doctrine of organicism are:

- *The centrality of the organism as the starting point of biological theorizing.* Just as the concept of energy occupies a central role in physics, the concept of an organism is the unit from which biological concepts and laws derive.
- *The importance of organization, as opposed to matter, as the defining characteristic of living organisms.* Organicists tend to eschew references to entelechies and substances in favor of emphasizing large-scale collective effects.
- *The autonomy of biology with respect to the physical sciences.* Organisms are more than complex machines in virtue of their distinctive kinds of organization. For example, growth is more than predetermined genetic programming; it involves intricate timing and ordering. Biological understanding of organic life requires concepts and laws that are particular to biology and have no counterpart in physics.

Vitalism and Evolution

Vitalism in relation to evolutionary theory is motivated by considering the role of organisms and their immediate surroundings in an explanation for adaptive design. After the early 20th century discovery of Gregor Mendel's (1822–1884) work on genetics, evolutionists tended to view the relation between development and evolution as unimportant for the investigation of the processes of natural selection. Yet, 21st-century research on the role of interactions between developing systems and the various surrounding signals and forces in the production of adaptive structures have revitalized interest on the contribution that organisms make to the processes of evolution. Organisms are not merely passive objects at the nexus of internal (genetic) and external forces. Organisms do not merely respond to physical conditions of their environment; rather, they play an active role. As biologist and philosopher Richard Lewontin (1929–2021) argued, organisms and environments are co-determined or co-constructed. Organisms contribute to evolution in a variety of ways. They actively select features of their surroundings which are relevant to their own survival, and in doing so, they change

their surroundings in ways that affect other organisms. This new emphasis on organisms provides a new way of investigating inheritance, development, and evolution and has, among other things, launched the field of evolutionary, ecological development or *evo-eco-devo*.

Philosopher of biology Denis Walsh has extended Lewontin's view by arguing that organisms are goal-directed agents. Hence, the proper scientific investigation of organisms and their role in the processes of natural selection requires understanding the way that organisms initiate and receive changes in forces and signals. This is best accomplished by adopting a methodological vitalism and embracing a suit of concepts and theories that are often shunned by biologists including *goals*, *affordances*, *salience*, and *teleology*. Many of the features of Walsh's vitalism are encapsulated in the tenets of organicism, as described above.

The Gaia Hypothesis

The Gaia hypothesis is the hypothesis developed by the independent scientist James Lovelock (1919–2022), according to which the Earth should not be considered just the environment where life appeared and developed but should be taken as a fully integrated with the dynamics of the biosphere, such that life, in a relative sense, would be fully and systematically integrated with the planet as a whole. The Gaia hypothesis appeared explicitly when Lovelock and the biologist Lynn Margulis presented evidence for the idea that, since life appeared on Earth, the atmosphere has always been in homeostasis with life, even in periods of change. Far from being the result of mere chance, they argued, this is explained by the hypothesis that, when life appeared on the planet, it took control of the environmental conditions that would favor the further development of life. They concluded that the Earth's atmosphere should be considered as a component of the biosphere, instead of just the environment that provided conditions for the development of life.

In the following decades, Lovelock explored the Gaia hypothesis in detail, stating that the Earth should be considered to be alive, and providing evidence for and against this idea. This

hypothesis suffered criticisms and modifications over the years but is discussed even nowadays as a plausible account in ecology, atmospheric sciences, and evolutionary biology. As Lovelock himself and others have argued, self-regulation of Earth doesn't imply some sort of teleology (though initial formulations of the hypothesis had left this open to interpretations) and is compatible with natural selection as we understand it nowadays. Still, the Gaia hypothesis includes the thesis that life cannot be isolated and described just by stating the biochemical conditions within organisms; instead, life should be seen as fully integrated and harmonized with external forces, dynamics, and conditions. In this sense, as observers have pointed out, the Gaia hypothesis resembles the form of criticism that vitalists addressed to some mechanistic accounts of life, as addressed in the previous section of this entry.

André Ariew and Gesiel da Silva

See also Biology, Evolutionary; Inference to the Best Explanation; Laws of Nature; Laws (Scientific); Philosophy of Science; Scientific Realism; Systems Science

Further Readings

- Bechtel, W., & Richardson, R. C. (1998). Vitalism. In *Routledge encyclopedia of philosophy*. <https://doi.org/10.4324/9780415249126Q1091>
- Bergson, H. (1944). *Creative evolution*. Random House. (Originally published in French, 1907)
- Bognon-Küs, C., Chen, B., & Wolff, C. (2018). Metaphysics, function and the engineering of life: The problem of vitalism. *Kairos: Journal of Philosophy & Science*, 20, 112–140.
- Driesch, H. (1914). *The history and theory of vitalism*. Macmillan.
- Gilbert, S., & Epel, D. (2008). *Ecological developmental biology: Integrating epigenetics, medicine, and evolution*. Oxford University Press.
- Johnstone, J. (1914). *The philosophy of biology*. Cambridge University Press.
- Lewontin, R. (2002). *The triple helix*. Harvard University Press.
- Lovelock, J. (2000). *Gaia: A new look at life on Earth*. Oxford University Press.

- Lovelock, J. E., & Margulis, L. (1974). Atmospheric homeostasis by and for the biosphere: The Gaia hypothesis. *Tellus*, 26(12), 2–9. <https://doi.org/10.3402/tellusa.v26i12.9731>
- Maienschein, J. (1991). The origins of Entwicklungsmechanik. In S. F. Gilbert (Ed.), *A conceptual history of modern embryology* (pp. 43–61). Springer US.
- Nicholson, D. J., & Gawne, R. (2015). Neither logical empiricism nor vitalism, but organicism: What the philosophy of biology was. *History and Philosophy of the Life Sciences*, 37(4). <https://doi.org/10.1007/s4065601500857>
- Normandin, S., & Wolfe, C. T. (Eds.). (2013). *Vitalism and the scientific image in post-Enlightenment life science, 1800–2010*. Springer.
- Walsh, D. M. (2015). *Organisms, agency, and evolution*. Cambridge University Press. <https://doi.org/10.1017/CBO9781316402719>

W

WARRANT

The concept of *warrant* is a contemporary notion adopted by some philosophers—among whom Alvin Plantinga stands out—that derives from the classical concept of *justification*. Both warrant and justification are concepts used in epistemology to describe and evaluate why a subject *s* holds a belief *p*. In that sense, warrant and justification are intertwined notions that sometimes are used interchangeably and sometimes are used as two different notions; this makes it difficult to establish a clear dividing line between the two concepts. Perhaps the less controversial difference between warrant and justification has to do with the differences between the externalist and internalist accounts in analytical epistemology: that is, the difference in what types of entities or phenomena can justify a belief. The notion of justification is typically internalist because it reserves the justification of a belief to what is in the subject's perspective of the world (i.e., internal states of the subject, such as other beliefs, serve to determine when a belief *p* is justified). In contrast, the concept of warrant is typically externalist because it holds that the justification of a belief is determined by its relation to the external state of affairs (i.e., external sources can be used to justify a belief *p*).

Within a general framework of controversies about the meanings of both warrant and justification, Plantinga raised a general criticism of the

very idea of justification that, according to his vision, is shared by the different expressions of internalism (and by some philosophers akin to coherentism, reliabilism, and even externalism) that dominate British and American epistemology. Contemporary epistemologists, such as the *classical* Roderick Chisholm and Richard Feldman as well as William Alston, Laurence Bonjour, Alvin Goldman, and John Pollock, present a great deal of confusion concerning what the connection is between justification and rationality, justification and knowledge, and justification and evidence. The main argument is that these different accounts of justification constitute varieties of the thesis according to which knowledge (i.e., what it is to know that *p*) can be analyzed conceptually as justified true belief (i.e., *s* is justified in believing that *p* is true). In contrast, Plantinga proposed that justified true belief is neither necessary nor nearly sufficient for knowledge. Hence, warrant is whatever precisely makes the difference between knowledge and justified true belief.

Warrant, the Gettier Problem, and the Proper Function Condition

The notion of warrant is connected to the question of the epistemic value of a justified true belief. In this vein, some philosophers—such as Michael Bergmann and Tyler Burge—present the notion of warrant as a project of responding to the philosophical challenges that the Gettier problem has implied regarding the definition of the idea of

knowledge. In 1963, Edmund Gettier famously demonstrated that a belief may be justified and true but still fail to be knowledge (i.e., a justified true belief is necessary but not a sufficient condition for knowledge), which called into question the ancient Greek philosopher Plato's idea of the definition of knowledge. Since then, Gettier's three-page article has stimulated a renewed effort in philosophy to clarify what knowledge constitutes.

Hence, the difference between a justified true belief and knowledge is that a belief can turn out to be justified and true by *epistemic luck*—that is, by error—and therefore not be knowledge. Consider one of the two cases that Gettier raised in his seminal article, which has been followed extensively by similar cases and analyzed by other philosophers, such as Roderick M. Chisholm, Alvin Goldman, and Brian Skyrms. Suppose that Smith and Jones have applied for a certain job, and suppose that Smith has strong evidence for the following conjunctive proposition *p*: Jones is the man who will get the job, and Jones has 10 coins in his pocket. Therefore, Smith concludes (justifiably, by the transitivity rule of the identity) the following proposition *q*: the man who will get the job has 10 coins in his pocket. In the end, Jones does not get the job; Smith gets the job. In addition, upon opening his pocket, Smith discovers he has 10 coins in it. Consequently, his belief *q*, according to which “the man who gets the job has 10 coins in his pocket,” was justified and true. However, Smith's belief turns out to be true by epistemic luck and therefore is not knowledge.

The Gettier cases demonstrate that a belief can be justified and true (in the sense of corresponding to the facts) even though the epistemic operation that supports it, in most cases where it would be used similarly, would lead to error. However, the formal process does not catch this error. These types of cases have two points in common. First, justification is fallible. In other words, the belief is justified and true, but this is not a sufficient condition for knowledge. Second, chance intervenes in the formal process of justification. These two points are combined in such a way that epistemic luck *compensates* for the weakness of justification and erroneously transforms into knowledge a justified true belief that was not in reality knowledge.

Plantinga suggested that the prevention of this potential error depends on the correct use of all subject's intellectual faculties (what he called the *proper function condition*). Only the correct use of all subject's intellectual faculties can be the *warrant* that the result of subject's epistemic procedures generates knowledge and not merely true belief. Thus, defined as a proper function condition, warrant is that which allows one to distinguish between knowledge and justified true belief. In other words, Plantinga presented the idea of warrant as a condition to solve the famous Gettier problem. Any theory that fails to consider the proper function condition as necessary for warrant is subject to Gettier-style counterexamples.

Warrant, Justification, Internalism, and Externalism

The notions that justification and warrant share both evoke the idea of the epistemic value of a particular belief—that is, if, when, and how one can give a certain type of support to a particular belief (i.e., to have a sufficient warrant or justification to hold a particular belief). However, the concepts' definitions depend on the different epistemological approaches in which they are used. In this sense, following the differences between internalism and externalism, what in general terms separates the notion of justification from that of warrant refers to what types of factors can act to provide epistemic support to a belief. Internalists use the term *justification* more to maintain that a belief's epistemic value is determined by a subject's internal states or reasons, whereas externalists use the term *warrant* more to maintain that a belief's epistemic value includes facts external to a subject's reasons.

The thesis of internalism can be formulated as follows: The internal states of the subject bearing the belief are what determine whether the belief is justified. It is common for internalists to resort to two premises to support this idea: (a) a belief is justified only if one has good reasons for accepting it, and (b) only the subject's internal states (e.g., other beliefs or other mental states) can be considered reasons to believe (or to accept a belief). In contrast, externalism maintains that the facts determining a belief's epistemic value embrace external conditions, such as whether the belief is

caused by external sources that make the belief true. In this vein, externalism emphasizes states of affairs that are external to the subject. In other words, externalists tie warrant to properties external to mental activity.

Burge proposed a currently highly influential variety of the externalist notion of warrant as *perceptual entitlement*. Perceptual entitlement constitutes a type of epistemic warrant that, unlike justification, can account for the intuitive fact that certain creatures are often warranted to hold perceptual beliefs, even when they lack reasons for them or do not have the concepts needed to reason regarding them. Such creatures may be entitled to believe because they possess highly reliable perceptual systems that, although incapable of producing reason, produce nonconceptual representational states providing warrant for their perceptual beliefs. According to Burge, epistemically unsophisticated creatures such as prelinguistic children and higher nonhuman animals represent paradigmatic examples in this regard.

Warrant, Justification, Foundationalism, and Coherentism

If the distinction between internalist and externalist conceptions of epistemic values is relative to the *nature* of the epistemic justification (what types of entities or phenomena can justify a belief: internal justification vs. external warrant), then there is another classification of theories about epistemic values that involves the *structure* of justification. Depending on how the structure of justification is conceived, one can distinguish between foundationalism and coherentism, among other approaches.

Generally associated with the idea of warrant, *foundationalism* leads to the argument according to which if a belief is justified, then it is either a basic or noninferentially justified belief (it receives no justification from another belief) or an inferentially justified belief (it receives justification from another belief). In contrast, *coherentism*, generally associated with the idea of justification, leads to the argument that there are no basic or noninferentially justified beliefs, so if justified, then inferences from other beliefs would justify every belief.

Thus, foundationalism defends the existence of two types of beliefs according to their justification: that is, basic or noninferentially justified

beliefs and derived or inferentially justified beliefs. In this vein, certain interpretations of the idea of warrant—such as Burge’s notion of perceptual entitlement—play the role of non-inferentially justified beliefs (i.e., beliefs not requiring justification). In contrast, coherentism maintains a *doxastic monism*: that is, the thesis according to which all beliefs are of the same type regarding their justification, and all beliefs are inferential (i.e., all beliefs receive their justification from other beliefs).

Hence, foundationalism maintains an architectural conception of the theory of epistemic values. Justification is seen as an asymmetric and linear structure. The set of beliefs is conceived as a *building* whose foundations are noninferentially justified beliefs on which the other beliefs (i.e., derived beliefs or inferentially justified beliefs) are based. Basic beliefs do not require the same sort of justification and function as the *warrant* of inferentially justified beliefs. Basic beliefs do not receive their justification from any other belief, either because they are self-justifying (given their nature) or because they receive their justification from other kinds of entities or phenomena (i.e., nonbeliefs), such as perceptual experience.

Unlike foundationalism, which characterizes the theory of justification as an *edifice* in which certain beliefs are not required to be justified and in which such beliefs function as a warrant for inferentially justified beliefs, *coherentism* characterizes the theory of justification as a *network*. According to this conception of justification, one’s system of beliefs is a *lattice* structure: that is, a network of multiple and mutual relationships among beliefs in such a way that this connection and not any specific beliefs in a unidirectional way (i.e., the foundational or non-inferentially justified beliefs in foundationalism) serves as a justification for any of the system’s beliefs. A belief is justified if it is part of a coherent set of beliefs and to the extent that its inclusion increases the coherence of that set of beliefs.

Coherentists’ main criticism of foundationalism posits that there are no satisfactory asymmetric and linear justifying chains, and an arbitrary interruption of the chain of reason occurs. In contrast, foundationalists argue there are beliefs that no longer need justification, and this is the role that the idea of warrant occupies in

foundationalism. Foundationalists' main criticism of coherentism affirms that coherentism is unsuccessful because it sees warrant as involving only the relation among beliefs and does not consider that the nexus between external environment and belief is also crucial to warrant.

Rodolfo Sarsfield

See also Epistemology; Fact Versus Theory; Justification; Philosophy of Science

Further Readings

- Alston, W. (1989). *Epistemic justification*. Cornell University Press.
- Armstrong, D. M. (1973). *Belief, truth and knowledge*. Cambridge University Press.
- Burge, T. (2010). *Origins of objectivity*. Clarendon.
- Chisholm, R. M. (1982). *Perceiving: A philosophical study*. Cornell University Press.
- Goldman, A. I. (1979). What is justified belief? In G. S. Pappas (Ed.), *Philosophical studies series in philosophy: Vol. 17. Justification and knowledge*. Springer. https://doi.org/10.1007/978-94-009-9493-5_1
- Plantinga, A. (1993a). *Warrant and proper function*. Oxford University Press.
- Plantinga, A. (1993b). *Warrant: The current debate*. Oxford University Press.
- Sellars, W. (1997). *Empiricism and the philosophy of mind*. Harvard University Press.

WORLDVIEW

Worldview may be defined as a pretheoretical framework of belief that offers individuals and societies an account both of what the world is and what the world should be. The term functions as a calque, or borrowed translation, from the German term *Weltanschauung*. Thus, *Welt*, or world, and, *Anschauung*, or perception, results in an anglicized rendering as *world-and-life vision* or, more briefly, *worldview*. Established usage in the United States is *worldview*; the *Oxford English Dictionary* hyphenates the term as *world-view*, although some scholars in Commonwealth countries follow the widespread practice of omitting hyphens in compound words. Hence, both

worldview and *world-view* are legitimate renderings of *Weltanschauung*.

The Salience of Worldview

Academic inquiry in late modernity has witnessed an exponential increase in the sheer number of theories, many claiming to explain the same phenomena, often in radically conflicting ways. To take but one example, the field of communication studies, which arose amid public speaking instruction and media effects studies in the mid-20th century, has spawned numerous theories of communication, explaining the phenomenon in terms of symbolic construction, interpersonal interaction, information systems, semiotic shifts, media effects, social influence, ideological change, and many other dynamics too numerous to list here. Few of these theorists are able to agree even on the definition of the phenomenon they supposedly share in common: One disciplinary expert, Stephen Littlejohn, notes some 126 different definitions of communication. Other disciplines experience similarly centripetal spreading of theoretical developments.

Coping with this profuse changefulness has tempted scholars to seek an all-encompassing account—or what philosophers would call a univocal account—for why 10 theorists can address the same phenomenon with 10 distinct but apparently viable theories. Given the elusiveness of such an account, however, other scholars have instead attempted to explain cross-disciplinary and intra-disciplinary differences by appealing to a given theorist's intuitive conceptual framework or worldview. Because this concept mediates between deeply held beliefs and empirical experience, it can enable productive exchange among scholars whose ideas might otherwise get lost in intertheoretical translation.

The Structure of Worldview

Expositors of the worldview concept generally begin by turning to the first major theorist of worldview, Wilhelm Dilthey (1833–1911), a German philosopher and psychologist. He grounds his account of worldview in what he calls life itself. Although life is lived out in innumerable various ways and is notoriously resistant to analysis, it

nonetheless exhibits some commonalities across the human species. One such commonality is that the person cannot help experiencing the surrounding world as something more than a set of causal, material forces. This inescapable sense provokes reflection, especially about the vanity of life, the changeableness of the world, the intense pleasure of everyday life, and the inexorability of death. Dilthey insists that everyone experiences these dynamics of life differently, but all must reflect on them as the largest questions of existence. Such reflection gels, as it were, into particular ways of knowing the world.

These ways of knowing are prior to empirical inquiry or theoretical investigation. To be sure, careful theoretical reflection can abstract the basic beliefs that constitute a particular way of seeing the world. But the sources for worldview beliefs are not wholly determinate. For example, a person experiences the people that populate her everyday life as a relatively stable matrix of relationships. A sociologist could theorize about each of this person's relations in empirical terms: who enjoys more or less social influence, who is likely to succeed or struggle economically, what needs or expectations determine their well-being. But no such analysis can univocally account for the lived experience of these relationships, because such experiences are themselves constituent aspects of life as lived by people down the ages. No theorist can evade these experiences or the perspective they engender, because they function presuppositionally before empirical method and theoretical inquiry get down to work.

Dilthey makes clear that a worldview's basic presuppositions constantly confront larger realities than theory-makers usually attend to. In all the bewilderments of life, questions about where the person originates, how she matures, and where she is destined are constantly at the center of the human's deepest reflection. In response to especially terrifying realities of life and death, individuals and societies have always developed shared worldviews that console, explain, or animate human endeavor, whether they have found themselves in an ancient Egyptian temple or a Baptist pulpit or a college library.

If worldviews are pretheoretical, can anything be theorized about their formation or structure? Dilthey would say yes. Worldviews are formed, he

explains, as we struggle to discern how to comport ourselves in response to experience. These attitudes toward the world, which Dilthey classifies in terms of optimism or pessimism, make up the basic materials of religion and literature, mythology, and philosophy. If we study how these frameworks of belief take shape, we can spot certain stable patterns of development. First, the person, as a perceiving subject, will develop a vision of reality, or what Dilthey calls a cosmic picture. Within this picture, she watches the events, the objects, and the people of her own life. Her perceptions gradually help her establish meaningful connections between all these things. These connections then form a platform for the second stratum of a worldview, the basic emotional evaluations of lived experience. Does she find something disgusting? Does she find a person delightful? Does she find an experience boring? In each case, these judgments will create a platform for the third layer of her worldview: her behavior in the world. Guided by her cognitive vision of reality, animated by her emotional responses to experience, she will seek a rightness and wholeness in the middle of her relationships. Gradually, layer upon psychological layer, from cognitive vision to impassioned evaluation to volitional determination, the person will build an ever more encompassing view of the world framed by beliefs that she holds to be universally reliable. A mature worldview will enable her to seek to bring her life into line with the Good as she understands it. For Dilthey, then, worldview becomes a regenerative concept—*generative* in the sense of giving shape to life, *regenerative* in the sense of allowing for innovative shaping of life.

Dilthey's account of worldview focuses on the interiority of the world-viewer. This preoccupation with the inner life of the world-viewer made sense within Dilthey's milieu of German Idealism and Romanticism that prevailed from the late 1700s to the mid-1800s. But it may also have given short shrift to more materialist dimensions of worldview formation. Friedrich Nietzsche's (1844–1900) materialist critique reworked the concept of *Weltanschauung* in a way quite opposed to Dilthey's commitment to truth: In short, Nietzsche treated worldview as nothing more than a consequence of power moves and structures. One contemporary theorist of worldview who

takes seriously both Nietzsche's critique and Dilthey's account of worldview is the contemporary Canadian philosopher James Olthuis. In his essay "On Worldviews," he lays out a Diltheyan exposition of the structure of *Weltanschauung* but does so with a phenomenology that takes into account the problems of materiality and ideology raised by Nietzschean critique.

Whereas Dilthey might be said to use the worldview concept to explain differences that arise between and among philosophical theories, Olthuis instead focuses on differences that arise between worldviews themselves. He traces the evolution of these differences from the idealism of worldview's first expositors to the materialist and historicist expositions of theorists like Nietzsche, Karl Marx, and Sigmund Freud and on to conceptual developments made by linguists, philosophers of science, sociobiologists, and religionists. For each of these groups, Olthuis notes a tendency to elevate one element—politics, economics, sexuality, language, faith, and genetics—as the determining formative factor for *Weltanschauung*. In contrast with such reductiveness, Olthuis's phenomenological theorizing proposes that worldview emerges among all these factors as they shape and are shaped by each other.

The problem for worldview, as the concept has developed through late modernity, is how to reconcile its aspirations for universality and its tendencies toward perspectivalism. Olthuis proposes to deal with this problem by treating the concept as a mediating entity between a person's deepest beliefs about the world and her diverse experiences of that world. The mediation that worldview performs is bidirectional: belief shapes experience, just as experience shapes belief. This allows Olthuis's account of worldview to take seriously both people's sense that their worldviews are dealing responsibly with the biggest questions of existence and critical theorists' sense that people are often unaware of how determined their beliefs are by cultural or biological forces. On the one hand, people deal with big questions by committing themselves to large answers. On the other, people must acknowledge the indispensable influence of traditions, societies, educational systems, not to mention physiological and psychological conditions. In other words, they cannot get behind these things to some pure, interior identity whose

beliefs are unaffected by cultural forces. But their worldview, says Olthuis, can mediate between these two dimensions of human life—belief and experience—toward the creation of a more integral personality.

This mediation between the beliefs to which we entrust ourselves and the cultural conditions in which we live makes for an account of worldviewing that is at once flexible and risky. Accordingly, Olthuis addresses what happens when worldviews fall apart. Sometimes people's anxieties create a divide between belief and experience, as when someone professes to believe in an ethic of compassion and yet practices an ethic of severity. Other times, people disguise efforts to maintain economic or political or institutional security by rationalizing their own power moves as ethically responsible. Consequently, worldview formation depends a great deal on both psychological well-being and social awareness. Olthuis acknowledges at times a worldview crisis can become so severe that people are forced to alter their deepest beliefs, though often in ways that they can recognize as maturation.

The History of Worldview

These accounts of the structure of worldview risk fixing a concept that has shown a remarkable flexibility in its short history. In contrast with the term *theory*, which traces back to an ancient Greek word (*theorein*) for the highest seat in a Greek theater, the worldview concept emerged only as late as the 1790s, when Immanuel Kant (1724–1804) apparently coined the term. In *Critique of Pure Judgment*, Kant's neologism *Weltanschauung* referred to the capacity of the human mind to intuitively apprehend the world. The new term became immensely popular in German intellectual culture by the 1830s, before spreading with memetic quickness across Europe.

Several factors seem likely to have contributed to the rapid popularity of the worldview concept in high modernity. For one thing, the term depends on an immensely influential ocular metaphor for describing epistemology. Of course, a commitment to visualist ways of knowing is hardly a modern invention. Indeed, the early Greek word for *theory* evokes a philosophical habit of looking steadily at something, thus detaching the theorist from the

changeable flow of everyday life and granting a platform for contemplating universal truths. This universalist emphasis, as Hjordis Becker notes, appears in Kant's development of the worldview term as a universalist concept for the knowing seer of the world, as well as in Dilthey's exposition of worldview in large, abstract typologies for religion, art, and philosophy. But although the term *worldview* continued the long-established habit of uniting knowing and seeing, Walter Ong has noted the ways that the rise of early modern science and the maturation of print technology altered and intensified the visualist tendencies of Western culture. Although Kant could coin the term *Weltanschauung*, and Dilthey could theorize it, neither could predict the semiotic shifting of the worldview concept, especially its shifting toward the perspective of the individual seer.

The newly invented term could hardly have enjoyed a more fortuitous historical moment for achieving currency. But the conditions that made it quickly popular also sped its evolution in conspicuously individualist and aesthetic directions. Charles Taylor has traced how these individualist dynamics found further support in notions of the interior self, cultivated by the introspective and egalitarian priesthood of all believers in the Protestant Reformation and extended by the Western tradition's conflicted support for two ideals at once: extraordinary heroism and ordinary life. These individualist forces, so supportive of the personalist vectors in the worldview concept, would achieve an eventual critical mass in the German Romantic movement of the 18th and 19th centuries, when thinkers like J. G. Fichte, F. W. J. Schelling, and Friedrich Schleiermacher repositioned the knower not as a theorist contemplating the inert glory of being but rather as a world-viewer making sense of the historical contingencies constitutive of everyday experience. As Albert Wolters notes, this repositioning represented a shift in attention from *ousia* (what is permanent and essential and universal) to *Geschichtlichkeit* (what is changeable and historically particular). This romanticization of epistemology enabled the worldview concept to develop not only in rational but also in aesthetic directions—directions still discernible today in many theorists' predilection for talking about worldview in terms of narrative and imagination.

The aestheticism of the worldview concept has from its origins created a tense relationship with the rationalist aspirations of theoretical inquiry. But perhaps no thinker put the work of world-viewing in tension with theory-making more vigorously than the Danish philosopher Søren Kierkegaard (1813–1855). In works such as *Either/Or* and *Fear and Trembling*, he depicted different ways of understanding life in terms of the aesthetic, ethical, and religious worldviews, or to use the term he preferred, *life views*. His literary criticism made clear that for him, one hugely important criterion for evaluating artistic effort was whether or not the artist helps people develop their own mature life views.

The worldview concept also caught the attention of religious thinkers, especially in those branches of Christianity emergent from the Protestant Reformation. Perhaps the most well known of religious worldview theorists was the Dutch prime minister and theologian Abraham Kuyper (1837–1920). He articulated the Reformed Christian faith in terms of a life system antithetical to other systems such as paganism, Islam, Roman Catholicism, and modernism. Kuyper's Scottish contemporary James Orr (1844–1913) also drew upon the concept of worldview, as an expression of the human yearning for a holistic account of the universe, in order to combat modernism. In the 20th century, philosophers like Herman Dooyeweerd (1894–1977) and Francis Schaeffer (1912–1984) helped to ensure that evangelical descendants of the Reformation would give more attention to the worldview concept than any other group of philosophers or theologians—an impressive distinction given the pervasive popularity of the worldview concept. As worldview historian David Naugle notes, *Weltanschauung* enjoyed a remarkable and virtually unparalleled flourishing in countries far from its first blooming and development.

The Uses of Worldview

Any exposition of the worldview concept in late modernity must address the fact that many intellectuals have appropriated the worldview concept under terms of their own devising. As an initial example, Karl Marx's (1818–1883) and Friedrich Engels's (1820–1895) notion of ideology showed

significant overlap with *Weltanschauung*, although Marxist philosophy distinguishes between the two by saying that ideology is more focused on socio-economic concerns than is worldview. This widespread reappropriation under other names raises the question of why the worldview concept has struck many people as useful, while the term itself has not. It may be that the scholars using the term have tended to do so in an institutional environment that privileges novelty and innovation: The publishing requirements of the academic guild have a built-in momentum toward neologism. Perhaps, too, the very popularity of the worldview term has resulted in a sloppiness that invites theorists to rename it for the sake of definitional precision. Furthermore, if, as some theorists claim, all theorists of worldview have a worldview themselves, they can hardly help redesigning terms that more closely align with their own worldviews. In what follows, this entry traces three different major appropriations of the worldview concept: phenomenological, scientific, and linguistic.

Phenomenological Appropriations

Keenly concerned about the seeming exhaustion of rationalism and idealism, Edmund Husserl (1859–1938) developed a new branch of philosophy called phenomenology. He argued that although disciplines like geometry and the physical sciences appear to be autonomous, they are actually planted in a pretheoretical world that nourishes those disciplines' significance. Once uprooted from this pretheoretical ground, even the most rigorous of disciplines wilts and collapses. In response to this exhaustion, Husserl developed a reflective mode of recovering the human experience of originary life in hopes of replanting theoretical inquiry in lived experience, where it could flourish. Thus described, Husserl's phenomenology sounds a great deal like a Diltheyan critique of naturalism. But Husserl was vigorously critical of the worldview concept for being contaminated by historicist and relativist tendencies.

At first, Husserl attempted to correct worldview by advancing what he called *Weltanschauung* philosophy. This philosophical enrichment of worldview arrives when careful critical reflection pursues wisdom and thus achieves a richer, fuller

development of what would otherwise be possible for the subjectivist perspective of the world-viewer. Husserl also tried to locate an even more rigorous mode of thought called *scientific philosophy*, which Naugle summarizes as the means for eliminating disagreement and worldview-ish relativism, as well as the establisher of foundations for scientific disciplines. But Husserl eventually despaired of this project, not least because Europe seemed to be overrun by personalist, subjectivist philosophies, which he attributed to the popularity of worldview. The danger, as he saw it, was not simply a loss of meaning through relativism; the danger was a loss of potential agency and action through cultural exhaustion. Husserl came to recommend that all human thought and practice be grounded in a recovery of an intuitive experience of the *Lebenswelt*, or lifeworld. This lifeworld arrives for the knowing subject, before she can superimpose theoretical categories upon it. The reflective person can regain both significance and strength by attending *Lebenswelt* as an enfolding dimension of experience prior to theoretical inquiry and rational conceptualization.

Are there commonalities between Husserl's lifeworld and the concept of worldview? Both concepts are pretheoretical. Both are attentive to the lived experience of the individual. Both have universal aspirations, though Husserl believed his notion of *Lebenswelt* could avoid charges of relativism better than worldview. Perhaps the most profound difference between the two concepts emerges in Husserl's profound distrust of history. In contrast with the more historically grounded notion of worldview, a lifeworld, at least in Husserl's telling, transcended all change and contingency, giving access to an enfolding, universal, even eternal, matrix for all human inquiry. The concept eventually came to exhibit more flexibility in the phenomenological tradition than Husserl granted it. Maurice Merleau-Ponty (1908–1961), for example, developed a historically informed *Lebenswelt* grounded in the human body. The importance of *Lebenswelt* as a kin concept to *Weltanschauung* for Husserl and Merleau-Ponty should not obscure the work of other phenomenologists, such as Karl Jaspers (1883–1969) and Martin Heidegger (1889–1976), who have addressed, at great length, the psychology and problems of worldview.

Scientific Appropriations

Despite the fact that worldview theorists have often treated scientific inquiry as a kind of foil for *Weltanschauung*, two scientists have achieved notoriety by drawing on the worldview concept in order to think through the philosophy and history of science: Michael Polanyi (1891–1976) and Thomas Kuhn (1922–1996). Polanyi's *Personal Knowledge*, published in 1958, appropriates worldview-ish thinking in his theory of tacit knowing. Arguing that we can know more than we can say, Polanyi insisted that what we know explicitly and theoretically is only a part of what we know tacitly and existentially. Consequently, we can only come to know reality by committing ourselves to living within a fallible worldview, or what Polanyi called an *interpretive framework*. What keeps such a framework from being solipsistic or relativistic? He wards off both concerns by counseling that the knower should toggle among three foci of attention: language, theory, and experience. Put differently, we shift our attention from the language we use to the theories we construct to the experience we live. In order to explain how this toggling enables indwelling an interpretive framework, Polanyi uses the simple but evocative example of a medical student who is trying to make a diagnosis using an X-ray film: At first, the student can see nothing on the X-ray but the patient's ribs; but as she listens to veteran radiologists talk about the X-ray, gradually she can come to see past the ribs to the pulmonary symptoms beneath. In other words, listening to the radiologist's *language* about their *experience* of the X-ray shapes their *theory* about the pulmonary condition. But what enables the student to make similar diagnoses is a tacit commitment to an indispensable interpretive framework. The radiology example might suggest that Polanyi's tacit frameworks bear only on limited, tightly framed problems. But he understands tacit knowledge to range from the smallest elements of empirical experience to the farthest reaches of intellectual and even religious inquiry. Polanyi would say that tacit commitments are indispensable for anyone trying to make sense of how language, theory, and experience bear on a vision of reality.

A much better known cousin to Polanyi's notion of interpretive framework emerged in

Thomas Kuhn's account of scientific discovery and problem-solving in his 1962 book *The Structure of Scientific Revolutions*. Part history of science, part philosophy of science, *Structure* argues for a sociology of scientific knowledge in which empirical inquiry pursues answers to communally agreed-upon problems within the ambit of what Kuhn calls a *paradigm*. The shared methods, instruments, beliefs, and problems that make up a given paradigm create the conditions of what Kuhn calls *normal science*. Within normal science, researchers agree on certain problems, make use of a common language (such as empirical formulae), and practice a shared comportment or attitude toward the problems before them. Sometimes, their paradigm proves unable to solve these problems or explain away various anomalies, which, if they become sufficiently numerous and persistent, create a crisis and bring about a scientific revolution. Researchers who embrace the resultant *abnormal science* will endure an untranslatability, or incommensurability, between the ways they used to see a given problem and the way their new paradigm requires them to see that same problem.

Kuhn used the term *paradigm* so diversely throughout *Structure* that even a sympathetic reader counted some 22 different meanings for the term. The terminological diffuseness of *paradigm*, however, may have made it especially suggestive for other fields and disciplines, many of which embraced the term in widely divergent ways. In a 1969 postscript to *Structure*, Kuhn himself tried to discipline his own semantics by delimiting two meanings for the term. His first sense of *paradigm* indicates what a community believes, values, and does in the practice of scientific inquiry (p. 175). This sense notices the community that forms among like-minded scientists and asks how we can explain what members of that community share in common that enables them to speak comprehensibly and to come to agreement when studying controversial phenomena. Kuhn's answer is they share a paradigm. Some scientists would prefer Kuhn to use the word *theory* instead of *paradigm*. To which he responds that that word, at least in its current usage, is too limited in scope for the holistic framework Kuhn is describing. (He later tried to switch out *paradigm* for *disciplinary matrix*. The new term never caught on.)

Kuhn's second sense of *paradigm* indicates the profile of a particular kind of problem, which, if grasped, enables scientists to solve other, apparently unrelated problems in analogous ways. This sense of the term counters conventional notions of scientific theory that assume that solving a particular problem requires only that one knows how to apply a pertinent theory with its concomitant rules. But Kuhn argues instead that the most important knowledge the scientists need for problem-solving exceeds both theory and rules. Scientists need to know not only the correct procedures for how to solve a problem but also a kind of transferable knowledge between problems. This knowledge bears less on cognition than on *praxis*: The scientist needs to spot analogies between apparently dissimilar problems and then to engage these new problems with the same approach required by already mastered problems.

How comparable are Polanyi's and Kuhn's accounts of tacit frameworks and paradigms to the usage historically attributed to worldview? Both thinkers highlight the sociality of knowledge and its incommensurability for people outside a given scientific community. Both emphasize the importance of the tacit and the praxis-based for the practice of scientific inquiry. But although worldview and Polanyi's notion of an inarticulable interpretive framework address large metaphysical questions, Kuhn's notion of a paradigm does not. A group of scientists could work within a paradigmatic understanding of physics or chemistry without addressing questions about the origins or destination of being. Such questions are not, of course, unrelated to wave motion or covalent bonding, but scientists taking up with the problems deemed appropriate to a given paradigm are not necessarily equipped by that same paradigm to address larger metaphysical questions. A paradigm, of course, can exhibit worldview implications. Indeed, some readers of Kuhn worried that his paradigmatic account of science was metaphysically relativistic. He responded to this criticism by saying that paradigms are always based on empirical evidence—which brings up a second difference between worldviews and paradigms: The latter are evidentially based. Worldviews may of course be grounded in empirical evidence, but they may as often as not be formed by other sorts of inquiry

as well, such as spiritual practice, hermeneutical inquiry, or ethical reflection.

Linguistic Appropriations

Thanks to the so-called linguistic turn in philosophy, in which theorists such as Ludwig Wittgenstein and Donald Davidson theorized language as a constructor, not simply a reflector, of reality, many theorists in the 20th century have remarked upon the reciprocal relationship between language and worldview. These theorists might be placed along a spectrum from those who hold that language gives shape to worldview to those who hold that worldview collapses into language. Among those who see language as constructive of but distinguishable from worldview, the linguist and rhetorician Kenneth Burke (1897–1993) developed a concept called *terministic screens*, very much in keeping with *Weltanschauung*. Burke noticed that despite scientific accounts of language as largely descriptive, people's use of symbols actually did much more than describe. Language, he noted, does at times *reflect* the world, but it also *selects* parts of the world for the language user to attend to and *deflects* other parts of the world away from the language user's awareness. For example, a pharmaceutical website may inform a website reader about aspects of a drug's ingredients, clinical effects, and any interactions with other drugs, but not discuss the sociological implications of the drug's usage or the rhetorical aspects of its marketing. This omission need not be underhanded: directing attention to one thing and not to another is an essential aspect of symbolic action. Accordingly, when we select a terminology, we are choosing a way of seeing the world. Burke appropriated the Augustinian maxim that we believe in order to understand by saying that we entrust ourselves to a way of speaking in order to see what it enables us to see.

Burke believed that the world existed independently from terministic screens. Later appropriators of a linguistic account of *Weltanschauung* were not so sure. Their doubts encouraged the notion that worldview is a kind of epiphenomenon of linguistic exchange. Richard Rorty has argued that humans invariably make use of a terminology to argue on behalf of what they do and believe, to account for what their own lives might

mean, and to praise or condemn what they hold to be good and evil. Rorty refers to this way of speaking as a *final vocabulary* because the terminology is *final* in that it cannot be argued for apart from the resources of language. A holder of the vocabulary might use a given terminology to defend almost anything, except itself. Questioning a person's final vocabulary moves that person to one of two recourses: the use of force or an attempt at redescribing the terminology in some new or appealing way. Rorty contrasts two basic approaches to these final vocabularies: common sense and ironist. The commonsense vocabulary invites an un-self-conscious articulation of ultimate beliefs as if they were ultimate truths that any rational being should assent to. Ironists, in contrast, know talking in *this* way and not *that* way enables people to see things as either good or bad, with the result that no ironist can allow herself to be complacent about her current habits of talk.

Are the linguistic frames described above equivalent to worldviews? Linguistic appropriations of *Weltanschauung* have the strange double consequence of both emphasizing and decentering the subjectivity so integral to the worldview concept. Rorty's final vocabulary, more than Burke's concept of the terministic screen, emphasizes the importance of individual subjectivity and the private projects of the self. On the other hand, they also diffuse the notion of a private, interiorized, centered, essential self. Burke's account of terministic screens, while not denying the inner life of the language user, is less concerned with what a language user internally means by such and such a term than with how language is functioning in and around the speaker. Rorty's notion of final vocabulary does even more to decenter the subject, to the degree that the self is nothing more than a tissue of sentences. Terministic appropriation of *Weltanschauung* both powerfully privileges and strangely diffuses the self.

The Critiques of Worldview

Despite the relative youthfulness of the worldview concept, it has enjoyed a remarkable popularity among both academics and the lay public. But perhaps in part because of the many and diffuse ways the term has been appropriated, it has also

endured stiff critique. This section traces three basic appraisals of *Weltanschauung*, focused on what worldview suggests about theory, subjectivity, and discourse. These criticisms tend to focus less on the worldview concept as defined in this entry than on reductions that tend to emerge when theorists use it sloppily.

Questions About Theory

The worldview concept is all too easily reducible to a rational system of beliefs and ideas, especially in apologetic or educational contexts, where movement from an intuitive, pretheoretical understanding to a determinate, theoretical explication feels necessary. But such a reduction loses track of what Charles Taylor has identified as the lived experience that is always the foundation for one's understanding of the world. In his analysis of contemporary religious belief and unbelief, what seems to make late modernity so decidedly secular is not that so many people have embraced one monolithically secular worldview (the sort of univocally irreligious or scientific system that Kuyper or Husserl feared). Instead, people today find themselves living with *conditions of belief* that make many different kinds of secularism and religiosity apparently reasonable and respectable. Worldview constructed theoretically could easily miss the conditions that make belief or disbelief possible, even when successfully naming the ideas that constitute a given framework of belief.

Questions About the Subject

The idealistic and Romantic roots of *Weltanschauung* make it vulnerable to critiques of individualism. Worldview can of course be understood communally. But as the philosopher William Rowe has argued, worldview only makes sense if we first posit a prior subject. At times, this subjectivity is construed in rationalist terms, especially when worldview is understood in a more systematic and cognitive fashion. At other times, a worldview-ish subjectivity tracks with a more aesthetic, Romantic notion of the self. But in any case, *Weltanschauung* seems to necessitate a peculiarly modern anthropology founded in a centered, interiorized conception of the self. Students of premodern accounts of the self, as well as theorists

of postmodern anthropology, will recognize the limitations of this latent individualism within the worldview concept.

Questions About Discourse

Olthuis's account of worldview as a mediating concept (see under "The Structure of a Worldview") helps explain why the interaction of language (itself a medium) and *Weltanschauung* is so complex. Does language provide a matrix from which worldview emerges? (As Rorty would seem to say.) Or is worldview prior to language? (As some of Rorty's realist opponents would seem to say.) This complexity makes it easy for theorists to resort to reductionisms. Rhetorician Celeste Condit has traced academic assumptions that discourse and ideology are entirely exchangeable for each other. Again, as with the overly theorized or subjectivist accounts of worldview, such reductionism emerges most frequently when theorists attempt to use the worldview concept as analytic equipment. When a politician delivers a speech, critics can show how the discourse functions in tandem with the speaker's worldview. In such cases, the speech is sometimes assumed to express a worldview without remainder. But there are other cases, as Condit points out, when material and emotional dimensions in speech function in excess of the speaker's worldview. For example, a politician may, through powerful pathetic appeal, move an audience to an action that might otherwise contradict their belief system.

Even with these critical questions in mind, the worldview concept has distinguished itself as an elegantly useful one for theorists capable of addressing the need for both stability and universalism in a modern and late-modern world highly sensitive to contingency and subjectivity. The fact that it also demonstrates a conspicuous elasticity under a broad range of theoretical sobriquets suggests that some variation on the concept will be assisting theoretical and pretheoretical inquiry for a long time to come.

Craig E. Mattson

See also Knowledge; Metaphysics; Paradigm; Philosophy of Science

Further Readings

- Audi, R., Dilthey, W., & Husserl, E. (1995) *The Cambridge dictionary of philosophy* (2nd ed.). Cambridge University Press.
- Becker, H. (2011). From 'Weltanschauung' to 'Livs-Anksuelse': Kierkegaard's Existential Philosophy. *Humana Mente—Journal of Philosophical Studies*, 18, 1–18.
- Burke, K. (1968). *Language as symbolic action: Essays on life, literature, and method*. University of California Press.
- Condit, C. (2013). Pathos in criticism: Edwin Black's communism-as-cancer metaphor. *Quarterly Journal of Speech*, 99(1).
- Dilthey, W. (1957). *Dilthey's philosophy of existence* (W. Kluback & M. Weinbaum, Trans.). Vision Press.
- Dorfman, E. (2009). History of the lifeworld: From Husserl to Merleau-Ponty. *Philosophy Today*, 53(3).
- Kierkegaard, S. (1946). *A Kierkegaard anthology* (R. Bretall, Trans.). Princeton University Press.
- Kuhn, T. S. (1962). *The structure of scientific revolutions* (3rd ed.). University of Chicago.
- Kuyper, A. (1931). *Stone lectures*. Eerdmans.
- Littlejohn, S. W. (1999). *Theories of human communication* (6th ed.). Wadsworth.
- Naugle, D. K. (2002). *Worldview: The history of a concept*. Eerdmans.
- Olthuis, J. H. (1989). On worldviews. In P. A. Marshall, S. Griffioen, & R. J. Mouw (Eds.), *Stained glass: Worldviews and social science*. University Press of America.
- Ong, W. J. (1958). *Ramus, method, and the decay of dialogue: From the art of discourse to the art of reason*. University of Chicago.
- Polanyi, M. (1958). *Personal knowledge: Towards a post-critical philosophy*. University of Chicago Press.
- Rowe, W. V. (1989). Society after the subject, philosophy after the worldview. In P. A. Marshall, S. Griffioen, & R. J. Mouw (Eds.), *Stained glass: worldviews and social science*. University Press of American.
- Rorty, R. (1989). *Contingency, irony, and solidarity*. Cambridge University Press.
- Rorty, R. (1998). *Achieving our country: Leftist thought in twentieth-century America*. Harvard University Press.
- Taylor, C. (1989). *Sources of the self: The making of the modern identity*. Cambridge University Press.
- Taylor, C. (2007). *A secular age*. Belknap Press.
- Wolters, A. (1989). On the idea of worldview and its relation to philosophy. In P. A. Marshall, S. Griffioen, & R. J. Mouw (Eds.), *Stained glass: Worldviews and social science*. University Press of America.

Appendix: Resource Guide

Because the topic covered in the encyclopedia is not a standard academic discipline with its own journals and foundational textbooks, we have here elected to include an eclectic mix of resources for students encountering issues related to theory in STEM and wishing to explore them further. First are seminal books in the history, philosophy, and foundations of theory. This is followed by a list of journals broken into three parts more or less mirroring the organization of the encyclopedia: philosophy; formal disciplines in STEM; empirical disciplines in STEM. Finally is a list of websites for a range of STEM discipline organizations.

Books

- Aristotle (2012 [40 BCE]). *The Organon: The works of Aristotle on Logic Paperback*, Roger Bishop Jones (Editor), E M Edghill (Translator), A J Jenkinson (Translator), G R G Mure (Translator), W A Pickard-Cambridge (Translator). CreateSpace Independent Publishing Platform.
- Bacon, Francis (2000 [1620]). *The New Organon (Cambridge Texts in the History of Philosophy)*, Lisa Jardine (Editor), Michael Silverthorne (Editor). Cambridge: Cambridge University Press
- Boyle, Robert (2007 [1661]). *The Sceptical Chymist*, gutenber.org
- Cartwright, Nancy (1983). *How the Laws of Physics Lie*. New York: Oxford University Press.
- Cassirer, Ernst (1980 [1923]). *Substance and Function & Einstein's Theory of Relativity*. Mineola, New York: Dover.
- Einstein, Albert (1991 [1920]). *The Theory of Relativity*. Mineola, New York: Dover.
- Godfrey-Smith, Peter (2003). *Theory and Reality*. Chicago: Chicago University Press.
- Evelyn Fox Keller (1985). *Reflections on Gender and Science*. New Haven, Connecticut: Yale University Press.
- Feyerabend, Paul (1975). *Against Method*. New York: New Left Books.
- Friedman, Michael (2001). *The Dynamics of Reason*, Stanford: CSLI Publications.
- Haraway, Donna (1989). *Primate Visions: Gender, Race and Nature in the World of Modern Science*. Routledge.
- Harding, Sandra (1986). *The Science Question in Feminism*. Cornell University Press.
- Hempel, C. G. (1965). *Aspects of Scientific Explanation and other Essays in the Philosophy of Science*. New York: Free Press.
- Howson and Urbach (2006). *Scientific Reasoning: The Bayesian Approach* (3rd ed.). Chicago and La Salle, Illinois: Open Court.
- Hume, David (2007 [1739/40]). *A Treatise of Human Nature: Volume 1: Texts (The Clarendon Edition Of The Works Of David Hume) (Clarendon Hume Edition Series)*. Oxford: Clarendon.
- Kuhn, Thomas (1969). *The Structure of Scientific Revolutions*. Chicago University Press.
- Lakatos, Imre, and Alan Musgrave, eds. (1970). *Criticism and the Growth of Knowledge*. Cambridge, UK: Cambridge University Press.
- Lloyd, Elisabeth (1984). *The Structure and Confirmation of Evolutionary Theory*. Princeton, NJ: Princeton University Press.
- Peirce, Charles (2014). *Illustrations of the Logic of Science*. Chicago and La Salle Illinois: Open Court.

Poincaré, Henri (1952 [1905]). *Science and Hypothesis*, Mineola, New York: Dover.

Popper, Karl R. (2002 [1959]). *The Logic of Scientific Discovery*, New York: Routledge.

Sober, Elliott (1984). *The Nature of Selection*, Chicago: Chicago University Press.

Suppe, Frederick (1977). *The Structure of Scientific Theories*. Chicago: University of Illinois Press.

Suppes, Patrick (1993). *Models and Methods in the Philosophy of Science: Selected Essays*. Dordrecht: Kluwer Academic Publishers.

Toulmin, Stephen (2003). *The Uses of Argument*. Cambridge University Press.

van Fraassen, Bas C. (1980). *The Scientific Image*, New York: Oxford University Press.

Whewell, William (1847). *The Philosophy of the Inductive Sciences, Founded Upon Their History*, 2nd edition, in two volumes, London: John W. Parker.

Journals

Theory of Science

The British Journal for the Philosophy of Science
www.journals.uchicago.edu/toc/bjps/current

Erkenntnis
www.springer.com/journal/10670

European Journal for the Philosophy of Science
www.springer.com/journal/13194

International Studies in Philosophy of Science
www.tandfonline.com/journals/cisp20

Metascience
www.springer.com/journal/11016

Perspectives on Science
direct.mit.edu/posc

Philosophy of Science
www.cambridge.org/core/journals/philosophy-of-science

Studies in the History and Philosophy of Science
sciencedirect.com/journal/studies-in-history-and-philosophy-of-science

Synthese
<https://www.springer.com/journal/11229>

Formal Disciplines

Annals of Mathematics
<https://annals.math.princeton.edu/>

Biometrika
<https://academic.oup.com/biomet>

Journal of Mathematical Logic
<https://www.worldscientific.com/worldscinet/jml>

Journal of Symbolic Logic
<https://www.cambridge.org/core/journals/journal-of-symbolic-logic>

Nature Machine Intelligence
<https://www.nature.com/natmachintell/>

Publications Mathématiques de l'Institut des Hautes Etudes Scientifiques
<https://www.springer.com/journal/10240>

Review of Symbolic Logic
<https://www.cambridge.org/core/journals/review-of-symbolic-logic>

Empirical Disciplines

American Journal of Physics
<https://www.aapt.org/Publications/ajp.cfm>

Astronomy and Astrophysics
<https://www.aanda.org/>

Cell
<https://www.sciencedirect.com/journal/cell>

IEEE Spectrum
<https://spectrum.ieee.org/>

Journal of the American Chemical Society
<https://pubs.acs.org/journal/jacsat>

Journal of High Energy Physics
<https://www.springer.com/journal/13130>

Nature
nature.com

Nature Communications
<https://www.nature.com/ncomms/>

Neural Networks
<https://www.journals.elsevier.com/neural-networks>

Physical Review

<https://journals.aps.org/>

Physical Review Letters

<https://journals.aps.org/prl/>

Proceedings of the National Academy of Sciences

<https://www.pnas.org/>

Science

<https://www.science.org/>

The Astrophysical Journal

<https://iopscience.iop.org/journal/0004-637X>

The European Physical Journal

<https://www.epj.org/>

Technology in Society

<https://www.sciencedirect.com/journal/technology-in-society>

Trends in Cognitive Sciences

<https://www.cell.com/trends/cognitive-sciences/home>

Other Online Resources

American Association for the Advancement of Science (aaas.org)

American Association of Physics Teachers (aapt.org)

American Chemical Society (acs.org)

American Geophysical Union (americangeosciences.org)

American Physical Society (aps.org)

Institute of Electrical and Electronics Engineers (ieee.org)

International Society for Technology in Education (iste.org)

National Academies: Sciences, Engineering, Medicine (nationalacademies.org)

National Science Foundation (nsf.gov)

Nature (nature.com)

Index

- Abbe, Ernst, 713
Abduction, 1–10, 380, 906
 in everyday life, 1
 in law, 3
 in medical sciences, 2–3
 in philosophy of science, 3
 scientific example of, 4
 in statistics and applied mathematics, 3
 subjectivity of, 3–4
 in use, 1–3
Abductive reasoning, 2, 788
Abel, Niels Henrik, 501
Abelard, Peter, 589, 784
Absolute dating, 312
Absolute knowledge, 774
Absolute scarcity, 611
Absolute zero, 734
Abstraction, 7, 10–15
 ancient, 12
 in computer science, 14–15
 language and, 11–12
 in mathematics, 13–14
 natural and scientific, 12–13
 in psychology, 14
 theory of, 11
Abstract knowledge, 5–10
 abstract thinking and, 6–8
 developing, 9–10
 human world and, 5–6
 moral sense and, 8–9
Abstract thinking, 6–10
Abu Said al-Sijzi, 705
Academic language, 435
Accuracy, 15–22
 of computation, 17–18
 of measurements, 18–21
ACDM model, 200
Achinstein, Peter, 276–277
Ackermann, Fran, 412
Ackers, Gary, 93
Activity theory, 392
Act of Creation, The, 190
Actor-network theory (ANT), 22–24
 cognitive science and, 144
Actual infinity, 416
Acupuncture, 569
Adaptation, 260, 888
 environment of evolutionary adaptedness and, 261–262
 of infants to garner adult attention, 265
 to social life, 265–266
Adaptation climate policy, 142
Adaptive logics, 631
Adaptive systems, 888
Addison, Thomas, 573
Additivity test, 886
Adelard of Bath, 515
Adenosine triphosphate (ATP), 69, 120–121
Ad hoc hypothesis, 24–26
Adler, Alfred, 220, 284, 585, 790–791
Adorno, Theodor, 814
Advance organizers, 586
Aepinus, Franz, 697
Affect, 414–415
Affective neuroscience, 623, 777
Agent-based models (ABMs), 178, 302, 453
Age of Reason. *See* Enlightenment, the
Age of systems in medicine, 547–549
Aghion, Philippe, 45
Agricola, Georgius, 235
Agricultural experiments, 267–269
Aim and Structure of Physical Theory, The, 405
Airs, Waters, Places, 549
Airy, George Biddell, 713
Al-Balkhi, Abu Ma'shar, 705
Al-Battani, 705
Al-Biruni, Abu Rayhan, 514, 705
Alchemy, 125
Alcoholic fermentation, 64–66
Alcuin of York, 706
Al-Dīn al-Qazwīnī al-Kātibī, Najm, 705
Al-Din al-Shirazi, Qutb, 705
Alexander, Hubert Griggs, 161, 163
Al-Fakhri, 514
Algebra, 528

- Algebra, Boolean, 520–521
 Algebraic theory, 288
 of the Renaissance, 527–529
Algebra of Logic, 472
 Algorithms, 290
Algorithmus Demonstratus, 527
 Al-Haytham, 705
 Alignment in bioinformatics, 78–79
 Al-Karaji, Abu Bakr Muhammad ibn Hassan, 513–514
 Al-Khwarizmi, Muhammad ibn Musa, 513, 526, 528, 705
*Al-kitab al-mukhtasar fi hisab al-jabr wa'l muqabala (The
 Compendious Book on Calculation by Completion and
 Balancing)*, 513, 526
Allegory of the Cave, 773
 Allochems, 314
 Allomorphs, 462
 Allopathic medicine, 552
 Allopoietic systems, 889
Almagest, 527
 Alpha, 869
 Alpher, Ralph, 199
 Al-Razi, 705
 Alston, William, 945
 Alternative hypothesis, 869–870
 Alternative medicine, 576
 Al-Tusi, Nasir Eddin, 527
 Alvord, Lori Arviso, 165–166
 Amagat, Émile, 710
 American Medical Association (AMA), 574
 Amicable numbers, 510
 Amoeba, 118
 Amontons, Guillaume, 709
 Ampère, André-Mari, 177, 708
 Ampliative inference, 375–376
 Analogical thinking, 40
 Analogy, 366, 906
 Analysis, 26–31, 435, 510
 critiques of, 30
 deductive approaches to, 27–28
 inductive approaches to, 28
 qualitative, 29–30
 quantitative, 28–29
 Analysis of variance, 268
Analytical Mechanics, 768
 Analytic functions, 510
 Analytic generalizations, 306
 Analytic geometry, 507
 Analytic logic, 485
 Analytic number theory, 393
 Analytic philosophy, 30
 Analytic theory, 153–154
Anatomica, 560
*Anatomical Disquisition on the Motion of the Heart and
 Blood in Animals*, An, 561
Anatomical Procedures, 560
 Anatomy in Enlightenment medicine, 560–561
 Anaxagoras, 682
 Ancient abstractions, 12
 Anderson, Alan Ross, 301, 631
 Andrews, Thomas, 710
 Angiography, 626
 Angle trisection, 685
 Angstrom, Anders Jonas, 709–710
 Animation, scientific, 454–455
 Annotation, linguistic, 458
 Anthropocentrism, 242
 Anthropological linguistics, 459, 463
 Antibiotic resistance and stewardship, 372–373
 Antibodies, 75
 Antimode, 866
 Antiquity
 mathematics in, 501–506
 medicine in, 547–552
 physics in, 680–686
 reasoning in, 783–784
 Antirealism, 280, 780, 799–800, 916
 Antiscientism, 279
 Apollonius, 526
 A posteriori, 424–425, 782
 Apparatus, measurement, 18–19
 Appel, Kenneth, 539
 Applicability of mathematics in science, 31–36
 connections with other debates, 34
 facing the challenge in, 31–34
 indispensability arguments for mathematical Platonism
 and, 35
 mathematical explanation in science and, 34–35
 mathematical modeling in science and, 34
 philosophical challenge of, 31
 Applied linguistics, 459, 463
 Applied mathematics, abduction in, 3
 Appolonius, 529
 Appropriate technology movement, 239
 Approximate designs, 270
 Approximate truth, 16
 A priori J, 424–425, 472, 782
 Aquinas, Thomas, 435, 559, 589, 659, 706
 Arab world
 influence on mathematics during the Renaissance,
 526–527
 mathematicians of, 513–514
 medicine in the Middle Ages in, 567
 medicine in the Renaissance in, 582
 Arbitrariness in language, 460
Archaeology of Knowledge, The, 815
 Archetypes, 434
 Archimedes, 319, 526, 674, 680, 682
 Archimedes' Principle, 682

- Arendt, Hannah, 47
 Argand, Emile, 746
 Argand, Jean-Robert, 518
 Argument indicator words, 493
 Aristarchus of Samos, 682
 Aristotle, 6, 12, 86, 249, 534, 587, 682
 “all is number” doctrine, 32
 applicability issue and, 32
 biological taxonomy, 894
 categorical syllogism of, 217–219, 291, 496–497, 663
 classical theory of concepts, 181
 on deduction, 217, 787
 on dissection, 549
 on first philosophy, 587
 as founder of logic, 470–471, 483–484, 492
 Galileo on, 400–401
 geometry and, 318
 on induction, 363, 478
 on measurability of nonquantitative properties, 544
 mental models and, 583
 physical theory and, 675
 on rationality, 773
 realism of, 434
 on truth, 927
 on types of motion, 704
Arithmetic, 515
Arithmetica Integra, 528
Arithmetica Logarithmica, 531
 Arithmetic and algebraic studies of the Renaissance, 527–529
 Arithmetic average, 865–866
 Arminger, Richard, 601
 Armstrong, David, 443, 588
 Arrhenius, Svante August, 89
 Arrhenius’ theory, 89
Ars Magna, 528
 Art, abstract, 13
 Artifacts, 212
 Artificial general intelligence (AGI), 36
 Artificial intelligence (AI), 36–41
 ad hoc hypothesis and, 25
 artificial neural networks (ANN), 38
 beginnings of, in mathematics, 289
 big data and, 59–60
 complex systems and, 177–178
 creativity of, 41
 how to think and talk about, 41
 human mind and, 40–41
 paradigms of, 39–40
 symbolic expert systems, 37–38
 symbolic reasoning systems, 36–37
 thinking by, 38–39
 types of, 36–38
 Artificial neural networks (ANN), 38
 Artistic perspectives of cognitive science, 148
Art of Conjecture, The, 508
Art of Speculation, The, 851
 Asch, Solomon, 777
 Aselli, Gaspare, 561
 Asenjo, Florencio G., 630
 Ashby, Ross, 887
Assayer, The, 32, 400
 Assimilation, 468
 Assumption, 41–45
 defined, 42–43
 examples of, 43–44
 Asthenosphere, 311, 742–743
 Aston, Francis William, 127
 Astrobiophysics, 90, 96
Astronomia Nova, 530
Astronomical Treatise, An, 516
 Astronomy. *See* Cosmology
 Astrophysics, 195–196
 Attention Deficit Hyperactivity Disorder (ADHD), 254
 Aumann, Robert, 301
 Austin, John L., 817
 Authoritarianism, 46
 Authority, 45–48
 charismatic, 46
 legal-rational system of, 46
 obedience to, 47–48
 political, 45–46
 psychological, 46–47
 religious, 46
 scientific, 45
 traditional, 46
 Autopoiesis, 207
 Autopoietic group systems, 889
 Averroes, 559, 705–706
 Avery, Oswald, 73, 82
 Avicenna, 566–568
 Avogadro, Amedeo, 88
 Axelrod, Robert, 301–302
 Axiology, 249
 Axiomatic theory, 52–55, 770, 816
 applications in various sciences, 55
 constraints, links, and theory nets in, 54–55
 elements and models of, 53–54
 observational and theoretical statements in, 52–53
 Axiom of choice, 534
 Axiom of infinity, 534
 Axiom of reducibility, 534
 Axiom schema, 48–52
 in logic, 48–50
 in mathematics, 50–51
 Babbage, Charles, 524
 Bachelard, Gaston, 451

- Bachelier, Louis, 851
 Backward error, 18
 Backwards induction, 299
 Bacon, Francis, 26, 109, 364, 398, 401, 561, 773–774, 812
 induction and, 478–479
 physical theory and, 675
 reasoning and, 785
 Bacon, Roger, 706–707, 784
 Bad science, 762
 Baker, Alan, 35
 Baker, David, 96
 Ball, Robert, 712
 Ballard, Danny Ramsey, 349
 Banach, Stefan, 641
 Banach-Tarski paradox, 641
 Bandura, Albert, 436
 Bardeen, John, 686
 Bargaining problem, 299
 Barlow, Willam, 89
 Barnes, Barry, 836
 Barré de St. Venant, Jean Claude, 711
 Barrow, Isaac, 507
 Barthes, Roland, 303
 Bartholin, Thomas, 561
 Barwise, Jon, 786
 Basalt, 310
 Basel problem, 509
Basic Laws of Arithmetic, 534
 Basic level of categorization, 823
 Basic Local Alignment Search Tool (BLAST), 78–79
 Batens, Diderik, 631
 Bateson, Gregory, 152, 159–160, 177
 Bau, 467
 Bawendi, M., 604
 Bax, Adriaan, 93
 Bayes, Thomas, 851
 Bayes factors, 856–857
 Bayesian statistics, 356, 377–378, 855–858
 applications of, 857–858
 challenges to, 857
 hypothesis tests and Bayes factors, 856–857
 Bayes' rule, 753–754
Bayes' Theorem, 851
 Beadle, George, 82
Beautiful Mind, A, 853
Beautiful; Seed Counting or Root Extractions, The, 516
 Beccaria, Abbé Giovanni, 87
 Bechamp, Antoine, 333
 Becker, Hjördis, 951
 Becquerel, E., 709
 Becquerel, Henri, 709, 736
 Bednorz, Georg, 730
 Beer, Stafford, 177, 890
Begriffsschrift, 219, 293, 524
 Behavioral economics, 147, 852
 Behavioral neuroscience, 623
 Behaviorism, 225
Behende und hübsche Rechnung auff allen Kauffmanschaftten, 528
 Belcher, Angela, 605
 Belenky, Mary Field, 934
 Belief, justified true, 423, 433
 Belief base, 57
 Belief persistence, 58–59
 Belief revision, 25, 57–58
 belief persistence and, 58–59
 definitions in, 57
 theories explaining, 57–58
 Belief set, 57, 425–426
 Bell, Alexander Graham, 236, 709
 Bell, John, 719
 Bell-shaped curve, 867
 Bell's Theorem, 678
 Belnap, Nuel, 631
 Beloch, Margherita Piazzolo, 539
 Beltrami, Eugenio, 325, 328
 Benacerraf, Paul, 780–781
 Ben-David, Joseph, 937
 Bendy geometry, 523
 Benjamin, Walter, 814
 Bentham, Jeremy, 611
 Bentler, Peter, 601
 Bentler-Bonnett nonnormed fit index, 601
 Berger, Hans, 290
 Bergmann, Michael, 945
 Bergson, Henri, 411, 417, 940, 942
 Berkeley, George, 531, 653–654, 779
 Berlin, Isaiah, 297
 Berman, Helen M., 95
 Bernal, Jon, 91
 Bernal, Martin, 323
 Bernard, Claude, 559
 Bernoulli, Daniel, 693, 697, 711, 768
 Bernoulli, Jacob, 508–509, 852
 Bernoulli, Johann, 508–509
 Bertalanffy, Ludwig Von, 176, 885, 942
 Berthelot, Marcellin, 65
 Berthollet, Comte Claude Louis, 89
 Berzelius, Jön Jacob, 125, 128
 Best systems, 916
 Best-systems accounts, 447–448
 Betti, Enrico, 711
 Bhaskar, Roy, 803
 Bhāskara II, 516
 Bichat, Xavier, 940
 Biddell, George, 712
 Big bang, 197–198, 701, 781

- Big data, 59–63, 111, 367, 756–757
 complex systems and, 179
 data mining, 60, 367, 854
 defined, 59
 ethical considerations with, 61–62
 philosophical foundation of, 59–60
- Big O notation, 393
- Bijection, 521–522
- Bimodal distribution, 866
- Binary analogue logic, 151
- Binding theory, 883
- Binnig, Gerd, 604
- Biocentrism, 242–243
- Biochemical rocks, 314
- Biochemistry
 19th century, 63–68
 Anselme Payen and diastase, 64
 contribution of Eduard Buchner to, 65–66
 roots of, 63–64
 sugar metabolism and enzyme specificity in, 66
 Wilhelm Kühne and physiological chemistry
 and, 66–67
 20th century, 68–71
 enzymes in, 69
 major instrumental advances in, 70
 metabolism in, 69–70
 molecular genetics in, 70–71
 neurochemistry in, 71
 vitamins in, 70
 contemporary, 72–77
 brain in, 75–77
 cells in, 72–73
 drug development and technology, 74–75
 fundamental concepts of, 72
 nucleic acids in, 73–74
 protein structure and function in, 74
- Bioinformatics, 77–78
 alignment toward structure and function in, 78–79
 sophisticated analytic tools in, 77–78
- Biological linguistics, 459, 463
- Biological membranes, 94
- Biological sciences, abstract knowledge in, 5–6
- Biological taxonomy, 894–895
- Biology, evolutionary, 79–88
 beginnings of genomics-based theory of, 85–86
 descent with modification, natural selection, and *The Origin of Species* in, 80–81
 diverging views prior to Modern Synthesis in, 82–83
 genomics in, 84–85
 problems arising after the Modern Synthesis in, 83–84
 science of heredity in, 81–82
- Biomedical model, 102–103, 343, 556–557
- Biomedicine, 72
- Biomolecules, 72
- Biophysics
 19th century, 86–87
 computational biology and, 95–96
 developments in physical chemistry in, 88
 electrochemistry in, 87–88
 electromagnetic theory in, 89–90
 later developments in, 88–89
 origins of, 86–87
 x-ray crystallography in, 88–89
 20th century, 90–97
 astrobiophysics, 96
 biological membranes in, 94
 medical, 96
 molecular, 92–94
 nerve impulse in, 92
 notable inventions in, 91
 structural biology and, 94–95
 contemporary, 97–102
 evolved optimality under physical constraints in,
 100–101
 minimal mechanistic models and Occam's rug in, 99
 modeling cellular individuality, 100
 as physics of living systems, 98–99
 theory in, 97–98
- Biophysiological theories, 348
- Bio-politics, 777
- Biopolymers, 68
- Bio-power, 777
- Biopsychosocial model, 102–106, 343–344
 biomedical and, 102–103
 causal interactions and, 106
 deep theory in, 103
 dualism and reduction problem in, 105
 etiology, treatment, boundaries in, 103
 evidence base for, 104
 mixed reception of, 103–104
 practical implications of, 104–105
 the psychological as embodied agency and the social as
 resource and, 106
 regulatory control and, 105–106
- Biostatistical theory, 553
- Biostatistics, 107–112
 defined, 107
 foundation phase of, 110
 genealogy of, 108–111
 gestation phase of, 107, 109–110
 institutionalization phase of, 107, 110–111
 practice/profession distinction in, 108
 terminology of, 107
- Biot, Jean-Baptiste, 88
- Bioterrorism, 373–374
- Biringuccio, Vannoccio, 235
- Bisociation, 190
- Bitcoin, 853

- Bjorkman, Pamela J., 94
 Black, Fischer, 852
 Black, Joseph, 125, 694, 709
 Black, Max, 589
Black Athena: Afroasiatic Roots of Classical Civilization, 323
 Blackbody radiation, 677
 Black-boxing, 837
 Black box of the mind, 814
 Black Death, 332, 369, 568
 Blackett, Patrick M. S., 742
 Black-Scholes pricing model, 852
 Black swan theory, 752–753, 851–852
 Blanqui, Louis-Auguste, 196
 Blastema-exudation hypothesis, 564–565
 Blind copies, 257–258
 Bloch states, 729
 Block, Steven M., 95
 Bloom, Benjamin, 435, 932
Bloom's Taxonomy, 435, 932
 Bloor, David, 836
 Blumenbach, Johann Friedrich, 563
 Bochar, Dmitrii A., 630
 Body politic, 550
 Boerhaave, Hermann, 548, 564–565
 Boethius, 514–515
 Bogdanov, Alexander, 176
 Bogen, James, 209–212
 Bohm, David, 720
 Bohr, Niels, 678, 689, 716, 737, 937
 Boltzmann, Ludwig, 88, 709, 712, 731, 734
 Bolyai, János, 325–328, 519, 523–524
 unreasonable effectiveness argument and, 782
 Bolzano, Bernard, 521
 Bombelli, Rafael, 528
 BonJour, Laurence, 424, 945
Book of Calculations, 515
Book of Healing, The, 567
 Book of Nature, 687
Book of Optics, 706
Book of the Devil Valley Master, 683
Book of the Extraction of Hidden Waters, 514
 Boole, George, 219, 292–293, 472, 520–521, 786
 Boolean logic, 219, 292–293, 786
 and algebra, 520–521
 Boorse, Christopher, 553
 Borel, Émile, 299–300, 417
 Borge, Jorge Luis, 107
 Born, Max, 675, 716, 737
 Bose, Satyendra Nath, 679, 737
 Bose-Einstein statistics, 737
 Bosons, 679, 737
 Boulding, Kenneth, 885
 Boundary objects, 835–836
 Bounded emotionality, 777
 Bounded rationality, 777, 853
 Bourdieu, Pierre, 836, 838
 Boussinesq, Joseph Valentin, 711
 Bowles, Robert, 302–303
 Box, George, 269–270
 Boyer, Carl Benjamin, 182
 Boyle, Robert, 125, 838
 Boys, Charles Vernon, 712
 Bradford-Hill, Austin, 554
 Bradford Hill criteria, 334, 554
 Bradley, Francis Herbert, 918
 Bradwardine, Thomas, 707
 Brahe, Tycho, 506, 676, 695
 Brahmagupta, 513, 516
 Brain, biochemistry of, 75–77
 Brain state-mental state identity theory (BSMSI), 657
 Bravais, Auguste, 89
 Briggs, Henry, 531
 Brink, Hilla, 343
 Bromberger, Sylvain, 275
 Bronfenbrenner, Urie, 584
 Bronk, Detlev Wulf, 90–91
 Brouwer, Luitzen Egbertus Jan, 299, 336, 416–418, 420, 510
 Brown, James Robert, 839, 922, 925
 Brown, Robert, 708
 Browne, Thomas, 697
 Brownian motion, 708
 Brunel, Isambard Kingdom, 236
 Brunel, Marc Isambard, 236
 Bruner, Jerome, 435
 Brünger, Axel T., 95
 Bubble Game, 853
 Bubonic plague. *See* Black Death
 Buchanan, Bruce G., 37
 Buchner, Eduard, 63, 65–66, 69
 Bucket theory, 936
 Bueno, Otávio, 34
 Buffon, George-Louis Leclerc, 563
 Bühler, Karl, 157–158
 Bullard, Edward, 742
 Bundle theory, 589
 Bunge, Mario, 280, 760, 772, 799, 801–802, 805–806
 Bunsen, Robert, 126, 709
 Burge, Tyler, 945, 947
 Burial metamorphism, 315
 Buridan, Jean, 705, 707
 Burke, Kenneth, 954
 Buschoff, Hermann, 582
 Bush, Vannevar, 127
 Bustamante, Carlos Jose, 95
 Butterfly effect, 539
 Byzantine Empire, 705

- Cailletet, Louis Paul, 710
 Calahan, Edward A., 852
 Calculus, theory of, 508–509, 531
Calculus of Antinomies, 630
 Calculus of reason, 520
 Calculus of variations, 508
 Calculus ratiocinator, 219
 Callon, Michel, 22
 Campbell, Norman, 543
 Cannizzaro, Stanislao, 126
Canon of Medicine, The, 332, 567
 Cantor, Charles, 93–94
 Cantor, Georg, 336, 416–417, 521–522, 534
 set theory, 825–831
 Cantor's theorem, 522
Capital, 814
 Capitalism
 accumulation of natural resources and, 610–611
 environmental studies and, 243
 growth of industrial, 610
 Carbohydrates, 68, 73
 Carbon sinks, 141
 Cardano, Girolamo, 850
 Cardano, Heironimo, 528
 Cardinality, 781, 825–826, 828–830
 Carnap, Rudolf, 52, 185, 280, 293, 366, 535, 633–634, 637, 817, 899, 914
 on induction, 363, 367
 linguistic frameworks and, 456–457
 theory, syntactic conception and, 917–918
 on theory and experience, 665
 Carnielli, Walter, 631
 Carnot, Sadi, 731
 Carnot function, 709
 Carret, Philip, 852
 Carroll, Patrick, 22
 Carrying capacity, 247
 Carson, Rachel, 128
 Cartan, Henri, 114
 Cartesian coordinate system, 506–507, 530
 Cartesian-split-mechanistic paradigm, 592
 Cartier, Jacques, 165
 Cartwright, Nancy, 916
 Cascade effects, 302
 Case-based reasoning (CBR), 179
 Case theory, 883
 Case-to-case generalizations, 306
 Caspar, Donald L., 93
 Cassinelli, C. W., 46
Castle of Knowledge, The, 528
 Casullo, Albert, 424
 Cataclysmic evolution, 83
 Categorical data, 864
 Categorical logic, 496–497
 Categorical models, 597
 Categorical syllogism, 217–219, 291
 Categories, descriptive, 295
Categories (Aristotle), 470
 Categorization, 823
 Category theory, 113–116
 categories, objects, and morphisms in, 114–115
 equivalence in, 115–116
 functors, 115
 history of, 114
 natural transformations, 115
 Cauchy, Augustin-Louis, 521, 711, 713, 768
 Cauchy's theorem, 521
 Causal condition, 780
 Causal/dispositional accounts, 448
 Causal explanations, 34–35
 Causal inference, 377–378
 Causal interactions, 106
 Causal models, 911–912
 Causal relations model, 890
 Causation, 554–555, 906
 Cavendish, Henry, 125, 693–694
 Cayley, Arthur, 521
Celestial Mechanics, 511, 693
 Cell theory, 116–122
 cell function, 118
 cell organization and size, 117–118
 cell types, 118–120
 cell wall, 120
 cytoplasm, 117, 120–121
 cytoskeleton, 122
 discoveries that led to, 117
 eukaryotic cells, 119–120
 modern, 117
 nucleus, 121–122
 origins and basic principles, 116–117
 plasma membrane, 120–121
 prokaryotic cells, 118–119
 viruses, 122
 Cellular automation, 453, 540
 Cellular individuality, 100
 Cellular neuroscience, 623
 Cellulose, 64
 Celsius, Anders, 578
 Center for Information Biology and DNA Databank of Japan, 78
 Central limit theorem (CLT), 511, 851, 869
 Central tendency, measures of, 864–866
 Chaffee, Steven, 185–186
 Chaiken, Shelly, 187
 Chandler, Cynthia Kay, 349
 Chang, C. C., 641
 Chaos theory, 175–176, 523
 Charismatic authority, 46

- Charpentier, Emmanuelle, 73
 Chase, Martha, 82
 Chasles, Michael, 712
 Chaves, Mark, 46
Chemical Abstracts, 123
 Chemistry, 122–129
 beginnings of, 123
 brief history of chemical practices, 125–129
 physiological, 66–67
 pluralism, theories, and models in, 123–125
 Cheselden, William, 562
 Chevreul, Michel-Eugène, 64
 Chi, 551
 China
 mathematics of ancient, 504–505
 mathematics of Middle Ages, 514
 medicine in ancient, 550–551
 medicine in the Middle Ages in, 568–569
 physics of ancient, 683
 yin and yang principles in, 550–551, 568–569
 Chip-cell interface, 605
 Chiropractic, 576
 Chisholm, Roderick, 945–946
 Chi-square test, 601
 Chittenden, Russell Henry, 67
 Chloroplasts, 121
 Cholera, 548
 Chomsky, Noam, 226, 306, 434, 460–461, 485, 585, 814, 883
 Christianity, 566–567
 Chromatin, 122
 Chromoplexy, 85
 Chromosomes, 122
 Chromothripsis, 85
 Chun Szi-sung, 582
 Church, Alonzo, 130, 204, 290, 631, 786
 Church-Turing thesis, 129–135, 420
 examples of how Turing machines compute functions, 133–134
 functions, 130
 significance of, 134–135
 term *effectively computable* in, 130–131
 terms *general recursive*, λ -definable, and *Turing Computable* and others, 131–133
 Chu Shih-chieh, 514
 Cieślak, Marek, 76
 Civil rights movements, 814–815
 Clairaut, Alexis, 712, 768
 Clapeyron, Benoit Paul, 709, 711, 731
 Clark, John A., 360
 Clarke, Adele, 835
 Classical mechanics, 676–677
 Classical statistics, 377
 Clause grammar and structure, 881–882
 Clausius, Rudolf Julius Emanuel, 88, 196, 710, 731
 Clebsch, Alfred, 711
 Climate change, 371–372
 observing, 138–139
 as social problem, 138
 Climate policy, 141–142
 Climate science, 136–143
 climate change as social problem and, 138
 climate policy and, 141–142
 contestation and skepticism around, 142–143
 foundational questions in, 136–137
 methods of, 138–141
 models of the climate in, 138–141
 observing climate change in, 138–139
 overview, 136
 Clinical linguistics, 459, 464
 Clinical neuroscience, 623
 Clinical treatments, 105
 Clinical trials, 269
 Clinic role in 19th century medicine, 571–574
 Clogg, Clifford, 601
 Closed systems, 173, 886
Clostridium difficile, 371
 Clustered regularly interspaced short palindromic repeats, 73
 Code of Hammurabi, 547
 Coenzyme A, 69
 Cogent inductive arguments, 494–495
 Cogito, 774
 Cognitive aesthetics, 148
 Cognitive computing, 149
 Cognitive cultural economics, 147
 Cognitive dissonance, 584, 933
 Cognitive economics, 147
 Cognitive ergonomics, 149
 Cognitive hypnotherapy, 148
 Cognitive management, 147
 Cognitive marketing, 147–148
 Cognitive neuroscience, 623
 Cognitive philosophy, 148
 Cognitive poetics, 148
 Cognitive pragmatics, 146
 Cognitive psychology, 148, 823
 Cognitive science, 143–151
 artistic and philosophical perspectives of, 148
 economic perspectives on, 146–148
 ethics and future of, 150
 explanation and, 277
 fuzzy-trace theory in, 144
 heuristics and schemas in, 144–145
 intercultural perspectives of, 148–149
 linguistic perspectives on, 145–146
 network and hybrid approaches to, 144
 neuroscientific perspectives of, 149

- psychological and sociological perspectives of, 148
 purpose in, 145
 research scope of, 145
 software engineering and, 844
 systemic approaches in, 144
 technological perspectives of, 149
 tools used in, 149–150
- Cognitive semantics, 146
- Cognitive sociology, 148
- Cognitive syntax, 146
- Cognitive theory, 814
- Cognitive zoology, 145
- Cognitivism, 813
- Cohen, Paul Joseph, 538
- Cohen's *d*, 357
- Coherence
- cohesion and, 880
 - with intuitions, 445
- Coherence theory, 58, 776, 927–928
- Coherentism, 249–250, 425, 947–948
- rationality and, 776
- Cohesion, 880
- Cohn, Emil, 736
- Coin flips, 209
- Colaço, David, 212
- Cole, Sydney W., 68
- Collins, Harry, 836–837
- Collins, John, 531
- Collusion, 299
- Columbus, Christopher, 223–224, 528, 810
- Colyvan, Mark, 34
- Combined standard uncertainty, 20
- Commitments in inferentialism, 385
- Common Core standards, 436
- Common pool resources, 609, 614
- Common sense, 591
- Communibiology, 194
- Communication models
- Claude Shannon, 158–159
 - Hubert Griggs Alexander, 161, 163
 - Karl Bühler, 157–158
 - mental models in, 585–586
 - Roman Jakobson, 160–161
 - Ruesch and Bateson, 159–160
- Communication repertoire, 346–347
- Communication theory, technical overview, 151–172
- compared to information theory, 155
 - human CT, 163
 - method and methodology, 153–163
 - theory and, 153–163
- Communication toolbox, 346–347
- Communicative meaning, 817–818
- Communism, 761, 936
- Company identity, 147
- Comparative fit index, 601
- Compendious Book on Calculation by Completion and Balancing, The*, 513, 526
- Complementary medicine, 164–169
- contemporary application of scientific method to, 167–168
 - prevalence in the United States, 166–167
 - types in the United States, 165–166
- Completeness, 169–172
- background on logic and consequence in, 169–170
 - deductive proof and, 170
 - entailment and, 170
 - of logical system, 171
 - relative consistency and, 795
 - and soundness in mathematical logic, 849
 - in statistics, 870–874
 - theory, 170–171
- Complex adaptive systems (CAS), 173, 175, 889
- Complexity of Cooperation, The*, 302
- Complex systems, 172–180, 887
- agent-based modeling (ABMs) and, 178
 - artificial intelligence and, 177–178
 - case-based reasoning (CBR) and, 179
 - chaos theory in, 175–176
 - complex adaptive systems (CAS), 173, 175
 - core concepts in, 174–175
 - cybernetics and, 177
 - data mining and, 179
 - dynamical systems and, 178
 - fuzzy logic and, 178–179
 - historic and current themes in, 175–179
 - system in, 173–174
 - systems theory and, 176
- Compositionality, 818
- Compositional meaning, 818
- Compositional word semantics, 818
- Computation, accuracy of, 17–18
- Computational biology, 95–96
- Computational linguistics, 458, 463
- Computational neuroscience, 623
- Computer experiments, 272–273
- Computer science. *See* Big data; Cybernetics, 20th century; Software engineering
- Computer simulations, 453–454
- Comte, Auguste, 250, 398, 403, 671–672
- Concept maps, 586
- Concepts, 180–185
- defined, 180–181
 - graphical organizations and, 184
 - implicit definitions and, 182–183
 - knowledge representations and, 182
 - meanings as, 818
 - and other scientific disciplines, 181–182

- in philosophy, 181
- and their relationships, 183
- Conceptual analysis, 185–189
 - concepts, constructs, and variables in, 185–186
 - dimensions and levels of measurement in, 187–188
 - process of explication in, 186–187
 - reliability and validity in, 188–189
- Conceptual blending, 190–192
 - fields of application of, 191
 - matrices and bisociation in, 190
 - mental spaces and cross-space mapping in, 190–191
- Conceptual framework, 294–296
- Conceptual knowledge, 435
- Conceptual thinking, 7
- Concise Introduction to Logic, A*, 217, 363, 366
- Concomitant method, 364–365
- Condensed matter physics, 727
- Condensed phase single-molecule fluorescence spectroscopy, 91
- Condit, Celeste, 956
- Conditional probability, 851
- Conditional syntactic structures, 882
- Confidence interval (CI), 350, 869
- Confidence level, 869
- Confidence limits, 869
- Conflict behavior, 301
- Conflict resolution, 59
- Conjecture, 792–793
- Conjectures and Refutations*, 809
- Conjoint measurement methods, 545
- Conjunction Fallacy, 476
- Connecting Humans to Emotions: A Reinterpretation of the Philosophy of Mathematics*, 323
- Conrad, Peter, 111
- Consequence and completeness, 169–170
- Conservation, 243–244, 769
- Consilience, 192–194
 - of academic cultures, 193–194
 - as scientific method, 192–193
- Consilience: The Unity of Knowledge*, 193
- Consistency in logical theory, 495
- Consistency proof, 795
- Constitutive values of science, 935–937
- Constraint accounts, 448–449
- Constructive empiricism, 405
- Constructive interference, 717
- Constructive type theory, 420
- Constructivism, 419, 553, 782
- Construct validity, 189
- Consumption (ailment), 570
- Contact metamorphism, 315
- Content internalism and externalism, 662
- Context of utterance, 816–818
- Contextualism, 592
- Continental drift theory, 741–746
 - constructivist perspectives on, 745–746
 - importance of, 746–747
 - rationalist perspective on, 744–745
- Continental idealism, 511
- Continental margins, 313
- Contingency in logical theory, 496
- Contingent falsehood, 495–496
- Contingent truth, 495–496
- Continent-continent collision, 313
- Continuity with science, 445
- Continuous data, 864
- Continuum hypothesis, 535
- Contraposition, 474
- Controlled vocabulary, taxonomy as, 894
- Control of variability, 268
- Control theory, 883
- Convenience sample, 254
- Conventionalism, 780
- Convergent margins, 313
- Conway, John Horton, 204, 540
- Cook, Stephen, 291, 540–541
- Cooker-cutter criticism, 27
- Cook-Levin Theorem, 541
- Cook's Theorem, 541
- Cooper, Leon, 686
- Cooperative learning groups, 933
- Cooperative principle, 463
- Coordinating definitions, 409
- Coordination, 880
- Copenhagen Interpretation, 690–691, 719
- Copernicus, Nicolaus, 4, 195, 316, 320, 529, 578, 722, 810–811
 - information theory and, 399–400
 - natural philosophy and, 560
 - physical theory and, 676
- Copi, Irving, 217, 363, 366, 368
- Core cultural issues, 347
- Corey, Robert Brainey, 92
- Coriolis, Gaspard-Gustave, 708, 768
- Correlation, 254–255
- Correns, Carl, 81
- Correspondence rules, 409
- Correspondence theory, 780, 927
- Cosmology, 4, 45, 194–201, 775
 - 20th-century observations in, 196
 - astrophysics and thermodynamics in, 195–196
 - big bang versus steady state, 197–198
 - consensus model and, 200
 - dark energy and other surprises in, 199–200
 - gravitational theory and, 701–702
 - Halley's comet and, 695–696
 - Newtonian universe and, 195
 - relativistic standard, 198–199

- during the Renaissance, 529–530
- static and expanding models of the universe, 197
- Coulomb, Charles-Augustin de, 693–694, 696–698, 711
- Coulomb's law, 444
- Coulson, Seana, 191
- Counterfactual accounts, 448
- Counter-norms, 832
- Coupled Model Intercomparison Project (CMIP6), 140–141
- Cournot, Augustin, 299–300
- Cours de Mathematiques*, 511–512
- Covering law of explanation, 667
- COVID-19, 103, 369, 857
- Cowles, Michael, 864
- Craig, Thomas, 712
- Craig, William, 918
- Crane, Robert K., 69
- Craver, Carl, 911
- Credible commitment, 301
- Crest of the Peacock: The Non-European Roots of Mathematics, The*, 323
- Crick, Francis, 73, 82, 93, 96, 940
- Critical commonsensism, 279
- Critical discourse analysis, 150
- Critical organizational theory, 777
- Critical rationalism, 801, 900
- Critical realism, 799, 803–804
- Critical theory, 154, 250–251, 813–814
- Critical value, 869
- Critique of Pure Judgment*, 774
- Critique of Pure Reason*, 654, 664, 812, 950
- Crooke, William, 709
- Cross-cultural child and adult language research, 486–488
- Cross-space mapping, 190–191
- Crowdsourcing, 854
- Crucial experiments, 837
- Crust, Earth's, 310, 315, 742–743
- Crystallography, 727–729
- Crystals, 727–729
- C-systems, 630
- Cuckholdry, 264
- Cullen, William, 548–549, 559, 564
- Culturally competent communication frameworks, 344–347
- Curie, Marie, 736
- Curie, Pierre, 736
- Curl, Robert, 128, 604
- Curry, Haskell, 204
- Cuvier, Georges, 79, 561
- Cybernetics, 20th century, 59, 177, 201–208, 436
 - autopoiesis and meaning of living form in, 207
 - distinction, awareness, and self-reference in, 204
 - Gödelian shift in, 206
 - key themes in, 201–202
 - language, reference, and self-reference in, 204–206
 - mathematics, language, and, 203–204
 - pioneers in, 202
 - second-order, 202–203
- Cybernetics: Or Control and Communication in the Animal and the Machine*, 177, 201–202
- Cytoplasm, 117, 120–121
- Cytoskeleton, 122
- Da Costa, Newton, 630
- D'Alembert, Jean, 693, 768
- Dalton, John, 88, 125
- Damasio, Antonio, 777
- Da Monte, Giambattista, 578
- Dance of agency, 23
- Daniel, Wayne, 108
- Darden, Lindley, 904
- Dark Ages, 559, 705
- Dark energy, 199–200
- Dark matter, 199–200, 702
- Darwin, Charles, 3, 80–82, 110, 221, 364, 763, 813, 916
 - on capacity and use of language, 5
 - Gaia theory and, 246
 - in historical context, 620–621
 - inductive reasoning and, 787
 - natural selection and, 260–261, 451, 616–617
 - thought experiments, 921, 923
 - values in science and, 939
- Darwin, Erasmus, 80
- Darwin, George Howard, 712
- Data and phenomena, 209–213
 - categorical, 864
 - in contemporary life sciences, 452–455
 - continuous, 864
 - data sets, 864
 - examples of, 209–210
 - explaining, 211
 - nature of phenomena, 212
 - noise in, 211–212
 - portability of data, locality of claims about phenomena, 210–211
 - restatement and defense of, 212–213
- Data curation, 215
- Data mining. *See* Big data
- Data models, 213–216
 - distinguishing raw data and, 214
 - evaluating, 215–216
 - ontology of data and, 213–214
 - production of, 215
 - purpose(s) of, 214–215
- Data structures, 290
- Data visualization, 215
- Davidson, Donald, 928
- Da Vigo, Giovanni, 580
- Da Vinci, Leonardo, 235, 529, 560, 810–811
- Davis, Martin, 538

- Davy, Humphrey, 125, 709
 Dawkins, Richard, 84, 300–301, 619
 Dear, Peter, 525
 De Bondt, Jacob, 582
 De Broglie, Louis, 737
 De Broglie-Bohm theory, 720
 Debtor field, 392
 Debye, Peter, 729
 De Chambaud, Méneuret, 941
De Circulie Magnitudine Inventa, 288
 Decision making
 models for, 295–296
 patient-centered communication and, 346
 Decision problem, 536
 Decision theory, 854
 Declarative statements, 486
Decline of the West, The, 690
De Cotagione et Contagiosis Morbis, 332
 Dedekind, Richard, 131, 524, 849
 Deduction, 217–224
 Aristotelian syllogism, 217–219
 completeness and, 170
 contemporary, 219–220
 covering law of explanation and, 667
 deductive-nomological model of, 221–222
 defined, 217
 hypothetico-deductivism, 222–224, 357–361
 philosophy of science and, 665–668
 Popper's falsification model of, 220, 286
 prediction and explanation and, 751–753
 scientific inquiry models, 220
 Deductive approaches to analysis, 27–28
 Deductive arguments, valid and invalid, 493–494
 Deductive indicators, 484
 Deductive inference, 375
 Deductive logic
 formal *vs.* informal, 473–477
 soundness in, 846–847
 See also Deduction
 Deductive-nomological (D-N) model, 221–222, 274–276, 368, 667, 751–752
 Deductive reasoning, 2, 292–293, 787, 905–906
 Deductive theory, 910
 Deductivism, 222
 Deep ecology, 248
 De Ferro, Scipione, 528
 Definitions, 498
 Deflationism, 928–929
 Degrees of relative acceptability, 883
De Humani Corporis Fabrica, 560–561
 De-idealization, 916
 Deism, 785
 De la Martinière, Germain, 562
 DeLano, Warren Lyford, 95
 Delian problem, 685
 De'Luzzi, Mondino, 560
De Magnete, 726
De Medicina Indorum, 582
 Democracy and pragmatism, 749–750
 Democritus, 317, 682
 De Moivre, Abraham, 851
 Demonstrative inference, 375–376
 De Morgan, Augustus, 476
De Motu, 767–768
 DENDRAL, 37
 De Nemore, Jordanus, 707
 Denotations, 819
 Deolalikar, Vinay, 541
De revolutionibus orbium coelestium libri VI (Six Books Concerning the Revolutions of the Heavenly Orbs), 529, 725
 Derived magnitudes, 545
 Desargues, Girard, 530
 De Sauvages, François Boissier, 558, 564
 Descartes, René, 321–322, 325, 434, 812
 Cartesian coordinate system, 506–507, 530
 classical mechanics and, 677
 in historical background of information theory, 398, 400–402
 legacy of, 512
 rationality and, 774
 rational mechanics and, 768, 770
 reasoning and, 785
 substance dualism and, 661
Descent of Man, The, 620
 Descent with Modification, 80–81
 Descriptive categories, 249, 295
 Descriptive geometry, 319–320
 Descriptive research, 295
 Descriptive statistics, 29, 864–867
 measures of central tendency, 864–866
 measures of dispersion, 866–867
 Desienhofer, Johann, 94
 Destructive interference, 717
Detailed Analysis of the Mathematical Rules in the Nine Chapters and his Fundamental Mutual Changes in Multiplication and Divisions, 514
Determination of the Coordinates of Places for the Correction of Distances Between Cities, 514
 Determinism, 250
 linguistic, 461
 Deutsch, Morton, 584
 De Vauban, Sebastien, 236
 Developmental neuroscience, 623
 Developmental systems theory (DST), 224–227
 definitions and characteristics of, 225–226
 future directions in, 227

- history of, 224–225
 key concepts and themes in, 226–227
Developmental systems theory: An Integrative Approach, 225
 Deville, Henri Sainte-Claire, 710
De Viribus electricitatis in motu musculari, 548
 De Vries, Hugo, 81–82
 Dewar, James, 710
 Dewey, John, 223, 295, 404, 434, 471, 747–749, 933
 Diachronic signifying language, 465
 Diachrony, 461
 Diagenesis, 314
 Diagonal argument, 522
 Dialectic logic, 492, 784
 Dialetheism, 498, 631
Dialogue Concerning the Two Chief World Systems, 32, 723
 Diastase, 64
 Di Bondone, Giotto, 529
 Diderot, Denis, 129
Die Axiome der Quantitat und die Lehre vom Mass, 545
Die Coss, 528
Die medizinische Physik, 87
 Differential calculus, 508
 Dilthey, Wilhelm, 948–950
 Din al-Tusi, 705
 Dipthongization, 468
 Dirac, Paul, 675, 716, 737
 Direct knowing, 411
 Direct realism, 781
 Dirichlet, Peter, 501, 712
 Dirty data, 216
 Dirty snowball model, 695
Discorsi, 922
 Discourse, 457
Discourse on the Method, 401, 774
Discourses and Mathematical Demonstrations Relating to Two New Sciences, 400
Discourses on Bodies in Water, 723
 Discreteness of language, 460
 Discrete state spaces, 431–432
 Discussive logic, 630
 Disease force, 563
 Disinterestedness, 761, 936
 Disjunction, 489–490
 Dispersion, measures of, 866–867
 Displacement in language, 460
 Dispositional essentialism, 443–445
Disquisitiones Arithmeticae, 288, 518
 Dissection, 549
Dissertatio de Arthritide: Mantissa Schematica: De Acupunctura: Et Orationes Tres, 582
 Dissimilation, 468
 Divergent plate margins, 313
 DNA (deoxyribonucleic acid), 73, 78, 82–86
 biophysics and, 93–94, 96
 horizontal transfer of, 84
 cell nucleus and, 121–122
 cell organization and size and, 117
 developmental systems theory (DST) and, 225
 Dobzhansky, Theodosius, 79
Doctrine of Chance, The, 851
 Dolostone, 314
 Dominant strategy, 299
 Dooyeweerd, Herman, 951
 Dopamine (DA), 76
 Doppler, Christian, 708
 Dörfler, Viktor, 411–412
 Double aspect theory, 661
 Double-bind situations, 850
 Double blinding, 269
 Doubling the cube, 539
 Doudna, Jennifer, 73
 Drake, Daniel, 549
 Drescher, Melvin, 300
 Drexler, K. Eric, 604
 Drezek, Rebekah, 605
 Driesch, Hans, 940–941
Drosophila melanogaster, 81
 Drucker, Peter, 436
 Drude, Paul, 729
 Drude model, 729
 Drude-Sommerfeld model, 729
 Drug development and technology, contemporary, 74–75
 Dualism, 105, 660–661, 780
 Duality in language, 460–461
 Dubin, Robert, 364
 Du Bois-Reymond, Emil, 87–88
 Dubrunfaut, Augustin-Pierre, 65
 Duhem, Pierre, 274, 282, 405, 710, 790
 Duhem-Quine thesis, 789–791, 902
 Dulong, Pierre-Louis, 709
 Dummett, Michael, 389
 Dunglison, Robley, 107
 Dunwoody, Sharon, 186–187
 Dupré, John, 592
 Durkheim, Émile, 359–360, 482
 Dutch tulip craze, 853
 Dutrochet, René, 87
 Dynamical systems, 178
 Dynamics, 711–712
 Dynamothermal metamorphism, 315
 Dyson, Freeman, 76
 Eagly, Alice, 187
 Earth, 309–316, 741–747
 absolute dating of rocks on, 312
 age of, 311–312

- composition of, 310
- geological structures of, 315
- Geologic Time Scale of, 311–312
- internal structure of, 311
- life on, 315–316
- plate tectonics of, 312–313
- relative dating of rocks on, 312
- rheology of, 310–311
- rocks and minerals on, 313–315
- structure of, 310–311
- Eastern Empire, 705–706
- Echo chamber, 59
- Eclectics, 576
- Ecocentrism, 242–243
- Ecological systems theory, 584
- Ecology. *See* Environmental studies; Natural resources
- Economic perspectives on cognitive science, 146–148
- Economic speculation, 852–853
- Eco-spirituality, 248
- Eddington, Arthur, 197, 284, 700, 791
- Eden, Colin, 411
- Edison Thomas, 236–237
- Edmonds, Dean, 38
- Educational linguistics, 459, 464
- Egan, Kieran, 933
- Egypt, ancient
 - mathematics in, 518–503
 - medicine in, 547–548
- Ehrlich, Paul, 575–576
- Eicher, James, 147
- Eichmann, Adolf, 47
- Eidetic science, 151
- Eidetic theories, 153
- Eilenberg, Samuel, 113–114
- Einstein, Albert, 33, 89, 194, 220, 222, 283–284, 535, 668, 719, 786, 790–791, 813, 937
 - bosons and, 679
 - on Brownian motion, 851
 - cosmology and, 197
 - critique of absolute simultaneity, 918
 - game theory and, 289
 - general theory of relativity, 325, 678, 688–689, 700, 739–740, 782
 - hole argument and, 702–703
 - language and content relativity theory of, 700
 - photoelectric effect and, 716
 - process of discovery by, 702–703
 - quantum physics and, 678, 737
 - special theory of relativity, 688–689, 738–739, 771
 - unreasonable effectiveness argument and, 782
- Either/Or*, 951
- Ekimov, A., 604
- Elasticity, 711
- Eldredge, Niles, 764–765
- Electric battery, 125
- Electrochemistry, 87–88
- Electroencephalography (EEG), 150, 626
- Electromagnetic force, 737
- Electromagnetic theory, 89–90, 676, 686
- Electron microscope, 92
- Electrons, 736
- Elements* (Euclid), 316, 318–325, 399, 505, 518, 529, 544–545
- Elements of Conics*, 529
- Elliptic geometry, 519
- Emden-Meyerhof-Parnas (EMP) pathway, 69–70
- Embedding, 880
- EMBL Nucleotide Sequence Database, 78
- Embodied cognition, 658–659
- Emergence in complex systems, 174–175
- Emergent properties, 803–804
- Empirical relativism, 837
- Empirical statements, 899
- Empirical theories, interpretation of, 407–410
- Empiricism. *See* Epistemology; Evidence; Experiment, theory of
- Empiricist structuralists, 876
- Empty sets, 825
- Empty space, geometries of, 322–323
- ENCODE Project, 78
- Energeia, 467
- Energy, 732–733
- Engel, George, 102–103, 344
- Engels, Friedrich, 196, 951
- Engineering
 - contemporary, 229–234
 - engineering theory and technology in, 231–232
 - engineers and the philosophy of technology, 232–233
 - theory in engineering, 233–234
 - theory in general, 229–231
 - history of, 235–241
 - early, 235
 - engineering systems, education, and professionalization in the late 1800s in, 236–237
 - European military engineering in, 235–236
 - Industrial Revolution engineering in, 236
 - late 20th- and early 21st-century developments and issues in, 238–240
 - through World War II, 237–238
- Engineering judgment, 230
- Engines of Creation*, 604
- Enlightenment, the, 812–813
 - See also* Renaissance, the
 - mathematics of, 506–513
 - medicine of, 558–565
 - physics of, 692–699
 - reasoning in, 785–786

- religion and, 698–699
 - social reform during, 699
- Enriques, Federigo, 329
- ENSEMBL, 78
- Entailment, 170, 486
- Entitlements in inferentialism, 385
- Entropy, 431, 731, 733–734, 887–888
- Enumerative induction, 478
- Environmental infectious diseases, 371–372
- Environmental justice, 247–248, 612
- Environmental studies, 241–248
 - anthropocentrism and, 242
 - capitalism and exploitation and, 243
 - carrying capacity and, 247
 - conservation and, 243–244
 - deep ecology and, 248
 - ecocentrism and, 242–243
 - environmental justice and, 247–248
 - ethnocentrism and, 242
 - Gaia theory and, 245–246
 - preservation and, 244
 - sustainability and, 246–247
 - Tragedy of the Commons and, 244–245
- Environment of evolutionary adaptedness, 261–262
- Enzymes, 64–67, 69, 74
- Epicurus, 364, 773
- Epidemiology, 370
- Epigeneticists, 224
- Epigenotype, 226
- Epiphenomenalists, 661
- Epistemic cognition, 433, 436
- Epistemic justification, 281–283
- Epistemic luck, 946
- Epistemological theses in scientific realism (SR), 805–806
- Epistemology, 248–252, 323, 433, 772
 - feminist, 252
 - foundationalism, coherentism, and reliabilism in, 250
 - historical and theoretical role of, 249
 - internalism and externalism and, 249–250
 - objectivist and subjectivist positions, 251
 - research paradigms and, 250–251
 - self-reflection and, 251–252
- Epsilon-delta definition of limit, 646
- Epstein, Ronald M., 344–346
- Equation of state, 732
- Equifinality, 586
- Equilibration, 584
- Equivalence, 495
 - object, 115–116
- Equivocation, 476
- Erasistratus, 550
- Erasmus of Rotterdam, 811
- Erdős, Paul, 289, 537–538
- Ergo clause, 774
- Ergon, 467
- Erikson, Eric, 933
- Erosion, 314
- Error argument, 806–807
- Error correction codes, 396
- Errors, 212
 - frequentist error probabilities, 861–862
 - in hypothesis testing, 354–356
 - measurement, 19
- ER transport network, 121
- Essay Concerning Human Understanding, An*, 774, 924
- Essay Towards Solving a Problem in the Doctrine of Chance*, 851
- Estimation, inferential claims for, 862
- Estimators, 867
- Eta-squared, 357
- Ether, 688
- Ethical data evaluation, 216
- Ethics*, 323
- Ethics and future of cognitive science, 150
- Ethnocentrism, 242
- Ethnolinguistics, 459, 464
- Ethnomethodology, 250–251
- Ethos, 427
- Euclid, 398–399, 505, 518, 529
 - classical geometry and, 316, 318–319
 - on measurability of nonqualitative properties, 544–545
 - non-Euclidean geometry and, 325
- Euclid Freed from Every Flaw*, 325–326
- Eugenics, 110, 621
- Eukaryotic cells, 119–120
- Euler, Leonhard, 509, 518, 540, 693, 711, 768
 - theory of classical mechanics, 677
- Euler's Circle, 509
- Euler's Identity*, 509
- Euler's Line, 509
- Europe
 - Black Death in, 332, 369, 568
 - mathematics in Middle Ages, 513–516
 - mathematics in Renaissance, 525–526
 - medicine in Enlightenment, 558–565
 - medicine in Middle Ages, 566–567
 - Scientific Renaissance in, 527
- European Bioinformatics Institute, 78
- European Organization for Nuclear Energy (CERN), 78
- Evaluation, 435
- Evaporites, 314
- Everett, Hugh, 720
- Everything Flows: Towards a Processual Philosophy of Biology*, 592
- Evidence, 252–260
 - contemporary theories of medical, 555–556
 - importance of assessment tools or outcomes as contributors to, 256–257

- increased emphasis on scientific, 253
- lack of versus negative, 258–259
- levels of scientific, 254–256
- in support of theory, 257–258
- weak, 253–254
- Evidence-based medicine (EBM), 555–556
- Evidence pyramid, 555
- Evidentialism, 776
- Evolution. *See* Darwin, Charles; Evolutionary psychology; Natural selection
- Evolutionary developmental biology (EDB), 227
- Evolutionary operation (EVOP), 270
- Evolutionary psychology, 260–266
 - adaptations of infants to garner adult attention in, 265
 - assumptions and principles of, 260–262
 - Darwin's theory of natural and sexual selection and, 260–261
 - developing fear of predators in, 262–263
 - environment of evolutionary adaptedness in, 261–262
 - evolved psychological mechanisms in, 261
 - investing in infants in, 264
 - jealousy in, 264
 - knowing what to eat and what not to eat in, 262
 - parental investment theory in, 263
 - parenting and, 264–265
 - reproduction and, 263–264
 - sex differences in the ideal mate in, 263–264
 - social relations and, 265–266
 - survival and, 262–263
- Evolution of Cooperation, The*, 301
- Evolved optimality under physical constraints, 100–101
- Evolved psychological mechanisms, 261
- Ewart, Peter, 708
- Ewing, Maurice, 742
- Excluded middle, the, 418–419
- Exclusion principle, 405
- Executive functions in mental models, 584–585
- Exercitatio Anatomica de Motu Cordis et Sanguinis in Animalibus (An Anatomical Exercise on the Motion of the Heart and Blood in Living Beings)*, 580
- Existential statements, 219
- Expectation states theory (EST), 913
- Experiential theory, 154
- Experiment, theory of, 267–273
 - George Box's industrial experiments and, 269–270
 - gravitational theory and, 700–701
 - Jack Kiefer and optimum experimental design, 270–271
 - robust designs, computer experiments, and, 272–273
 - Ronald Fisher and agricultural experiments and, 267–269
- Experimental theory, 154
- Experimenter's regress, 836–837
- Expert intuition, 411
- Explanation, 273–278
 - Carl Hempel's theory of, 274–276
 - cognitive science and, 277
 - contemporary theory of, 276
 - covering law of, 667
 - in early physical theory, 681
 - as form of communication, 276–277
 - lawlike and accidental generalizations, 275–276
 - mental models for, 277
 - naturalism and, 277
 - prediction and, 668, 751–753
 - probabilistic, 667
 - state of theory today, 278
 - truth and, 273–274
- Explanatory reasoning, 2
- Explanatory research, 296
- Explicandum, 221
- Explicans, 221
- Explication process, 186–187
- Explicit versus tacit reasoning, 788
- Exploratory research, 295
- Explorer's argument, 806
- Expression meaning, 816
- Expressive meaning, 822–823
- Extended mind theory, 659
- Externalism, 249–250, 425, 946–947
 - content, 662
 - rationality and, 776
- Externalities, 612
- Extrascientific values in science, 937–939
- Face of Earth, The*, 743
- Facets of intuiting, 413–414
- Facial electromyography, 150, 626
- Factorial designs, 268
- Factual knowledge, 435
- Fact versus theory, 279–283
 - fact, theory, and epistemic justification and, 281–283
 - ontological status of fact and epistemological status of theory and, 280–281
- Fahrenheit, Gabriel Daniel, 578
- FAIR data practices, 216
- Fallacies of Relevance, 476
- Fallacy, 498
- False conjectures, 792–793
- Falsifiability, 220, 283–287, 357–358, 365, 663, 760
- Falsificationism, 790
- Faraday, Michael, 708
- Farr, Wanda K., 64
- FASTA, 78
- Fast life history strategy, 261
- Fate, 25
- Fatou, Pierre, 289, 539
- Fauconnier, Gilles, 190

- Faunal succession, 312
Fear and Trembling, 951
 Fear of predators, 262–263
 Fechner, Gustave Theodor, 545
 Fedorov, Evgraf, 89
 Feedback, 888
 Feedforward, 888
 Feeding, 888
 Feigenbaum, Edward, 37
 Feigenbaum, Mitchell, 203
 Feldman, Richard, 945
 Felsic magma, 314
 Feminism, 349, 815
 Feminist epistemology, 252
 Fermat, Pierre de, 288, 510, 530, 850
 Fermat's Last Theorem, 539
 Fermentation, 64–66
 Fermi, Enrico, 675
 Fermi gas, 729
 Fernández, Facundo, 75
 Fernel, Jean, 578
 Ferrari, Lodovico, 528
 Fever, 548
 Feyerabend, Paul, 12, 671
 Feynman, Richard, 33, 604, 675, 678, 702
 Fibonacci, Leonardo Pisano, 515
 Fibonacci sequence, 515
 Fichte, J. G., 951
 Fick, Adolf, 87
 Fictionalism, 780
 Finite-element simulations, 453
 Finitely and irrevocably testable (FIT) theory, 919
 Finiteness theorem, 510
 Finite sets, 825
 Finitism, 416
First Lines on the Practice of Physics, 549, 564
 First-order logic, 535
 Fischer, Emil, 66, 69
 Fisher, Irving, 851
 Fisher, Ronald A., 81, 110–111, 350, 365, 368, 451
 agricultural experiments and, 267–269
 model-based frequentist inference and, 858–863
 Fitness landscapes, 887
 Fixed-point theorem, 299
 Fizeau, Hippolyte, 712
 Flagella, 118
 Flat curve, 868
 Fleck, Ludwik, 634, 833
 Flood, Merrill, 300
 Fludd, Robert, 581
 Focal point, 301
 Foliations, 315
 Folk beliefs, 591
 Folk psychology, 657
 Foolproof processes, 233
 Forbes, James David, 709
 Ford, Donald H., 225
 Ford, Henry, 237
 Forecasting
 prediction and, 757–758
 speculation as, 854
 Formal hypotheses, 296
 Formal inference, 376–377
 Formalism, 336, 474–475
 Formal logic. *See* Logic, formal and informal
 Formal sciences, 287–294
 as different from other sciences, 287
 foundations of logic in, 291–293
 foundations of modern-day mathematics in, 288–291
 importance of, 293–294
 Forman, Paul, 690
 Forman thesis, 690
 Formism, 592
 Forrester, Jay W., 889
 Forward error, 18
 Forward induction, 299
 Fossil fuels, 136
 Fossils, 261–262, 316
 punctuated equilibrium and, 763–766
 Fostering the patient-provider relationship, 345
 Foucault, Léon, 708, 712–713
 Foucault, Michel, 433, 777, 815, 929
 Foundationalism, 249–250, 425, 774, 947–948
 Foundation phase of biostatistics, 107, 110
Foundations of Behavioral Research, 185, 223
Foundations of Geometry, 323
 Foundation theory, 58
 Founder effect, 307
 4R model of neuroscientific research, 624–626
 Four-color theorem, 539
 Fourier, Joseph, 501, 511–512, 678
 Fourier series, 511
 Fowler, Ralph, 734
 Foxall, Gordon, 368
 Fracastoro, Girolamo, 332
 Fraenkel, Abraham, 538, 830
 Frame, John, 774
 Frames, 182
 Frameworks, 294–298
 conceptual, 294–296
 how to develop, 297
 linguistic, 456–460
 providing a lens, 296–297
 systems theory, 392
 types of research, 297
 Frank, Philip, 634
Frankfurt School, 814

- Frankl, Viktor, 59, 933
 Franklin, Allan, 283
 Franklin, Benjamin, 87, 693, 696
 Franklin, Rosalind, 73, 82, 93
 Franklin rod, 696
 Franz, Rudolph, 709
 Fraunhofer, Joseph von, 713
 Free choice sequence, 418
 Free-riding strategies, 302
 Frege, Gottlob, 33, 219–220, 288–290, 336, 417, 534, 849
 Begriffsschrift, 219, 293, 524
 on concepts, 181
 on logical connectors, 387
 mathematical logic and, 472
 non-Euclidean geometry and, 329
 Platonism and, 779
 predicate logic and, 497
 reasoning and, 779, 786
 third realm of, 779–780
 Freire, Paolo, 436
 Frensel, Augustine-Jean, 708
 Frequency, 866
 Frequentist statistics, 377
 Fresnel, Henri, 875
 Freud, Sigmund, 27, 220, 284, 434, 790–791, 933
 Frey, Gerhard, 540
 Freyd, Peter, 114
 Friedman, Alexander, 197
From Metaphysics to Ethics: A Defense of Conceptual Analysis, 185
 Fuchs, Stephan, 838
 Fu Hsi, 550
 Fujimura, Joan, 835
 Fuller, Buckminster, 583
 Fullerene, 604, 606
 Functionalism, 658
 Functional magnetic resonance imaging (fMRI), 149, 626–627
 Fundamental forces, 676–677
 Fundamental magnitudes, 545
Fundamental Mutual Changes in Multiplication and Divisions, 514
Fundamental Theorem of Algebra, The, 518
 Funk, Casimir, 70
 Fuzziness
 in syntactic structures, 882
 and variation in language, 467–468
 Fuzzy logic, 178–179
 Fuzzy-trace theory, 144
 Gaia theory, 245–246, 614–615, 943
 Galen, 552, 558–561, 566, 940
 Galilean transformations, 688
 Galileo Galilei, 32, 285, 316, 321, 398, 400–401, 433, 530, 578, 682, 695, 711, 722, 810–811, 900–901
 Accademia dei Lincei and, 563
 Age of Reason and, 506
 Book of Nature and, 687
 falling bodies thought experiment, 922
 legacy of, 723–724
 on nonphysical property measurement, 545
 physical theory and, 676
 rational mechanics and, 767–768
 reaction of the Church to, 723
 Galileo's paradox, 521
 Galison, Peter, 835
 Galois, Évariste, 517, 536
 Galois theory, 510
 Galton, Francis, 110, 620, 851
 Galvani, Luigi, 87, 548
 Gambler's Fallacy, 475–476
 Game theory, 289, 299–303, 853–854
 concepts in, 299–300
 continuing developments in, 302–303
 history of, 300–302
 Gamow, George, 198–199, 379
 Garbett, Samuel, 125
 Gardner, Howard, 435, 586, 932–933
 Gates, Bill, 239
 Gauss, Carl Friedrich, 268, 288, 417, 712, 851
 19th century mathematics and, 518–519
 constructing a regular heptadecagon, 685
 non-Euclidean geometry and, 325–328
 theory of calculus, 508
 Gaussian curvature, 519
 Gaussian distribution, 511, 519
 GenBank, 78
 General abstract nonsense, 113
 General circulation models (GCMs), 139–140
 General covariance, 700–701
 General Equivalence Theorem, 271
 Generality, 250
 Generalization, 7, 303–307, 478, 480
 logic of, 305
 in qualitative research, 306
 in quantitative research, 306
 symbolic, 634
 true, 440
 General relativity, 325, 678, 688–689, 700, 739–740
 unreasonable effectiveness argument and, 782
 General systems theory (GST), 885
 Generative grammar, 883
 Generativity of language, 460
Genetical Theory of Natural Selection, The, 451
 Genetic drift, 307–309
 distinction between selection and, 308–309
 varieties of, 307–308

- Genetic epistemology, 933
- Genomics, 84–85
- Genotype, 81
- Geochemistry, 310
- Geochronology, 310
- Geoffroy, E. F., 125
- Geographical linguistics, 459, 464
- Geologic Time Scale, 311–312
- Geology, 309–316
- age of the Earth, 311–312
 - geological structures, 315
 - life on Earth, 315–316
 - plate tectonics, 312–313
 - rocks and minerals, 313–315
 - structure of the Earth, 310–311
- Geometrical calculus, 507
- Geometry
- classical, 316–324, 684–686
 - Euclid's *Elements* and, 316, 318–319, 322, 399
 - extraterrestrial geometry and axioms in, 318
 - four dimensions of philosophy of geometry and, 323–324
 - geometric paradigms in, 319–320
 - geometries of empty space in, 322–323
 - geometries of the world and, 320–321
 - mechanical worldview in, 321–322
 - pre-Euclidean, 317
 - Pythagoras and, 317–318
 - during the Renaissance, 530
 - non-Euclidean, 324–331
 - acceptance of, 328–329
 - background and origins of, 325–326
 - discovery of, 326–328
 - implications for truth of mathematics and, 329–331
- Geometry, The*, 322
- Geometry of numbers, 524
- Geomorphology, 310
- Geophysics. *See* Geology
- Geothermal gradient, 310
- Gerard, Ralph, 885
- Gerber, Christophe, 604
- Gerhardt, Charles Frédéric, 126
- Germ theory, 331–335
- denialism, 333
 - Koch's postulates today and, 333–334
 - origins of, 331–333
- Gerstner, František Josef, 713
- Gesell, Arnold, 225
- Gestalt shifts, 834–835
- Gestation phase of biostatistics, 107, 109–110
- Gettier, Edmund, 249–250, 780, 946
- Gettier problems, 249–250, 437, 945–946
- Gherard of Cremona, 515
- Gibbs, Josiah Willard, 89, 710, 731
- Giere, Ronald, 817, 915
- Gilbert, William, 697, 722, 726
- Gilbreth, Lillian, 240
- Gintis, Herbert, 302–303
- Glashow, Sheldon, 737
- Gleick, James, 152
- Global Indigenous Data Alliance, 216
- Global north, 610
- Global south, 610
- Global surface temperature (GST), 137
- Glycolysis, 69
- Gödel, Kurt, 132, 220, 288, 330, 419–420, 535
- cybernetics and, 206
 - incompleteness theorem, 171, 534, 643, 795–796, 830–831, 849
 - on mathematical intuition, 786
 - reasoning and, 786
- Gödel's incompleteness theorems, 335–342, 534
- first and second, 336–337
 - historical background of, 335–336
 - reception of, 340–341
 - technical details of, 337–340
- God's sensorium, 325
- Goldman, Alvin, 945–946
- Goldschmidt, Richard, 83
- Golgi apparatus, 121
- Gooding, David, 921
- Good Judgment Project, 854
- Goodman, Nelson, 480
- Google, 436
- Gootsman, Eugene, 39
- Gorski, Philip S., 360
- Gosling, Ray, 82
- Gottlieb, Gilbert, 225
- Gould, Stephen Jay, 619, 764–765
- Government theory, 883
- G-protein-coupled receptors (GPCRs), 74–75, 95
- Grammar, 457
- clause, 881–882
 - transformational and generative, 883
 - universal, 461, 472, 477–478
- Grammatical change in language, 469
- Gramsci, Antonio, 252
- Grand antiprism, 540
- Grand theory, 815
- Granite, 310
- Graphene, 606
- Graphic organizations and concepts, 184
- Graphology, 457
- Graph theory, 393
- Grassmann, Hermann, 712
- Graunt, John, 112
- Gravitational mass, 676
- Gravitational paradox, 195

- Gravitons, 679
 Gray, Russell D., 226
 Great Internet Mersenne Prime Search, 540–541
 Greece and Rome, ancient
 See also Aristotle; Euclid
 mathematics in, 505–506, 525–526
 medicine in, 548–550
 physics in, 681–683
 Green, George, 711, 713
 Greenfeld, Liah, 937
 Greenhouse effect, 511–512
 Green party politics, 248
 Gregor, Shirley, 844
 Grelling, Kurt, 641
 Griesemer, James, 835
 Griffin, Cindy, 252
 Griffiths, Paul E., 226
 Grimm's law, 469
 Grim-trigger strategy, 299
 Groen, Guy, 919
 Grosseteste, Robert, 706–707
 Grossmann, Marcel, 330
 Grothendieck, Alexander, 114, 289, 537
 Grounded theory method, 28, 906
Grounde of Artes, The, 528
 Groupthink, 777
 Guggenheim, Ernest, 734
 Guigó, Roderic, 78
 Guldberg, Cato, 88, 710
 Gullberg, Jan, 827
 GUM approach, 18
 Gunnell, John, 222
 Gutenberg, Johannes, 721
 Gut feel, 412
 Guthrie, Francis, 539
 Gutiner, Johannes, 560

 Haack, Susan, 279
 Haber, Fritz, 127
 Habermas, Jürgen, 9, 776–777, 814
 on self-reflection, 251–252
 Habits, 58
 Hacking, Ian, 904
 Hadamard, Jacques, 524
 Haga, Kazuo, 539
 Hahnemann, Samuel, 576
 Haken, Wolfgang, 539
 Halas, Naomi, 605
 Haldane, J. B. S., 66, 81
 Hale, George Ellery, 714
 Hall, G. Stanley, 225
 Halldén, Sören, 630
 Halley, Edmond, 693, 695–696
 Halley's Comet, 528, 695–696

 Halliday, M. A. K., 880
 Halmos, Paul R., 827
 Halting function, 134
 Halting problem, 536
 Hamblin, C. L., 786
 Hamel, Georg, 770
 Hamilton, William Rowan, 521, 677, 708, 712, 768
 Hamlin, Christopher, 573
 Hannay, Jo Erskine, 844
 Hanson, Norwood Russell, 280, 283
 Hantaviruses, 371–372
 Haraway, Donna, 838
 Hardin, Garrett, 244–245, 613–614, 854
 Harding, Sandra, 838
 Hard systems (HSs), 173
 Hardy, Godfrey Harold, 451, 532–533
 Hardy-Weingberg principle, 451
 Harman, Gilbert, 380–381
 Harré, Rom, 478, 480–481
 Harsanyi, John C., 301
 Hartmann, Stephan, 454
 Hartshorne, Charles, 660
 Harvey, William, 560–561
 Hastie, Trevor, 597
 Häüy, René Just, 89
 Hawking, Stephen, 590–591, 675
Head Jewel of Accuracy, 516
 Heads or tails flip, 209
 Health and body, Enlightenment theories of, 558–560
 Health and diseases, contemporary theories of, 552–553
 Health care delivery, 343
 Health care science, 343–350
 health disparities and minority health in, 347–348
 models of health care delivery in, 343
 patient-centered communication and culturally competent communication frameworks in, 344–347
 women's health care in, 348–349
 Health disparities, 347–348
 Heat, 709, 732–733
 Heat pumps, 733
 Heawood, Percy, 539
 Hebb, Donald, 290
 Hebbian learning, 290–291
 Hebb's law, 291
 Hegel, Georg Wilhelm Friedrich, 30, 774, 812, 898, 918
 Hegemony, 252
 Heidegger, Martin, 952
 Heisenberg, Werner, 678, 690, 716, 737, 852
 Heisenberg's uncertainty principle, 852
 Helmholtz, Hermann von, 87–88, 574, 708–709, 712–713
Hè Megalè Syntaxis, 705

- Hempel, Carl, 220–222, 282, 368, 667, 751–752, 816–817
 induction and, 478, 480
 theory, syntactic conception and, 914, 918
 theory of explanation, 274–276
- Henri, Victor, 66
- Henry, Joseph, 708
- Heraclitus, 681, 773
- Herbal medicine, 567, 569
- Herbrand, Jacques, 132
- Herman, Robert, 199
- Herophilus, 550
- Herschel, John, 616, 709
- Herschel, William, 693
- Hershey, Alfred, 82
- Hesper, Ben, 77
- Hess, Harry Hammond, 742–744, 746
- Hessel, Johan, 89
- Hessen, Boris, 688
- Hessen thesis, 688
- Heuristics, 378–379, 412
 cognitive science, 144–145
- Heyting, Arend, 336, 419
- Hidden variables, 720–721
- Hierarchical universe, 195
- Hieroglyphs, 503
- Higher-level analysis, 29
- Higher order logic, 912
- Highspread, 866
- High-temperature superconductivity, 730
- Hilbert, David, 134, 323, 329, 336, 416, 716, 816, 914
 20th century mathematics and, 533–534
 axiomatics and, 770
 decision problem, 536
 reasoning and, 786
- Hilbert curves, 533
- Hilbert inequality, 533
- Hilbert's axioms, 511
- Hilbert space, 533, 716
- Hilbert's problems, 533
- Hilbert's program, 534
- Hilbert system, 48–50, 330–331
- Hill, Archibald, 92
- Hill coefficient, 92
- Hintikka, Jaakko, 919
- Hipparchus, 682–683, 705
- Hippocrates, 548–549, 552, 567
- Historical linguistics. *See* Linguistics, historical
- HIV/AIDS, 369
- Ho, W., 604
- Hobbes, Thomas, 36, 550, 838, 921
 rationality and, 774
 reasoning and, 786
 thought experiments, 923–924
- Hobson's choice, 850
- Hofweber, Thomas, 293
- Hogeweg, Paulien, 77
- Holder, Otto, 545
- Hole argument, 702–703
- Holland, John H., 175, 889
- Hollander, Alexander, 92
- Holling, Crawford Stanley, 173
- Holocaust, the, 47
- Homeopathic medicine, 576
- Homer, 398, 549
- Homogeneity test, 886
- Homological algebra, 537
- Homomorphisms, 115
- Homonymy, 821–822
- Hooke, Robert, 116–117, 711
- Hooke's law, 53
- Hopkins, Frederick Gowland, 68
- Hoppe-Seyler, Ernst, 63
- Horizontal DNA transfer, 84
- Horkheimer, Max, 814
- Horvath, Johann Baptiste, 693
- House of Wisdom*, 513
- Huang, Laura, 411, 413
- Hubble, Edwin, 196
- Hubble constant, 196
- Huber, Robert, 94
- Hull, David, 619, 940
- Hull, Gordon Ferrie, 709
- Hulse, Russell A., 701
- Human CT, 163
- Human Genome Project, 78, 557
- Human mind
 abstract knowledge and, 5–6
 artificial intelligence (AI) and, 40–41
- Humboldt, Wilhelm von, 467, 518
- Hume, David, 110, 180, 220–221, 325, 365
 bundle theory and, 589
 induction and, 479, 666, 792
 influence of Newtonian theory on, 688
 on laws of nature, 447
 phenomenalism and, 654
 rationality and, 776
 on reason as slave of passions, 471
 reasoning and, 786
 uniformity of nature assumption and, 753
- Humoral theory, 558–560
- Hunter, John, 562–563
- Hunter, William, 558
- Huntington, Edward, 219
- Hurley, Patrick, 217, 363, 366
- Husserl, Edmund, 152, 952
 Platonism and, 779
- Hutton, Charles, 545
- Huygens, Christiaan, 288, 713, 768

- Hybrid disciplines in linguistics, 463–464
 Hybridity in cognitive science, 144
 Hybrid theories of health and disease, 553
Hydrodynamica, 693
 Hydrodynamics, 712–713
Hydrodynamics, 768
 Hydrogeology, 310
 Hydrophobic effect, 74
 Hyperbolic geometry, 796
 Hypersocial species, humans as, 265
 Hypotheses, 906–908
 Hypothesis testing, 350–3435
 alternatives to, 356–357
 Bayes factors and, 856–857
 errors in, 354–356
 limitations of drawing conclusions from, 354
 logic of, 350–351
 one-tail versus two-tail, 352–354
 parametric and nonparametric test statistics in, 351–3430
 p value in, 352, 356–357
 stating the hypothesis in, 351
 Hypothetical syllogism, 217
 Hypothetico-deductivism, 222–224, 357–361
 critiques of, 358–359
 in the social sciences, 359–360
- Ibn al-Shatir, 705
 Ibn Hayyān, Jābir, 705
 Ibn Muhammad, Ja'far, 705
 Ibn Sina, 332
 Ideal epistemic communities, 761
 Ideal gas, 429–430
 randomness and probability in, 431
 Idealism and realism, 779–780
 Idealization, 471
 Ideals of natural order, 283
 Ideational function in language, 460
 Idols, mental, 774
 Idols of the cave, 478
 Idols of the marketplace, 478
 Idols of the tribe, 478
 If-clauses, 882
 Igneous rocks, 313–314
Iliad, 398, 549
 Illocution, 462
 Illocutionary theory of explanation, 276
 Imaginary geometry, 629
 Imaginary logic, 629
 Immediate inference, 472
 Implication, 495
 Implicit definitions and concepts, 182–183
 Incommensurability, 834–835, 903
 Incompleteness theorem, 220, 643, 830–831, 849
 Inconsistency in logical theory, 495
 Independent replication, 268
 Index minerals, 315
 India
 mathematicians of Middle Ages, 516
 medicine in ancient, 551
 physics in ancient, 683
 Indication value, 19
 Indirect measurement, 546
 Indirect realism, 781
 Indispensability, 781–782
 Individual psychology, 585
 Induction, 363–369
 consilience of, 192
 critique of, 479–481
 defined, 363–364
 early development of, 364–365
 justification of, 481–482
 in models of scientific inquiry, 367–369
 modern forms of, 366–367
 philosophy of science and, 665–668
 prediction and explanation and, 751–753
 probability and, 667–668
 problems of, 365–366
 refutation (R) and, 792–793
 in the social sciences, 482
 See also Logic, inductive
 Inductive approaches to analysis, 28
 Inductive arguments
 strong, weak, and cogent, 494–495
 valid and invalid, 493–494
 Inductive indicators, 484
 Inductive inference, 375
 Inductive logic. *See* Logic, inductive
 Inductive reasoning, 2, 292–293, 498, 773–774, 787, 905–906
 Inductive-statistical (I-S) model, 274–276, 368
 Industrial experiments, 269–270
 Industrial Revolution engineering, 236
 Infants
 adaptations of, to garner adult attention, 265
 investing in, 264
 Infectious disease studies, 369–374
 antibiotic resistance and stewardship, 372–373
 bioterrorism, 373–374
 environmental infectious diseases, 371–372
 infection prevention and control, 370
 vaccine research, 370
 Infectious heredity, 83
 Inference, 375–380, 817
 abductive, 788
 ancillarity and, 873–874
 categorizing, 375
 causal, 377–378

- deductive versus inductive, 375
- demonstrative versus ampliative, 375–376
- formal versus material, 376–377
- heuristics, 378–379
- immediate, 472
- statistical, 366–367, 377–378
- varieties of, 377–380
- Inference to the best explanation (IBE), 380–383
 - critiques of, 383
 - in different fields, 382–383
 - example of, 381–382
 - forming hypotheses using, 381
 - terminology of, 381
- Inferential conception, 34
- Inferentialism, 383–390, 878
 - applications in logic, 386–388
 - historical foundations of, 384–385
 - key concepts in, 385
 - linguistic rationalism and, 385–386
 - semantic theory and, 388–389
 - variations of, 389
- Inferential statistics, 29, 864, 867–868
- Infinite sets, 825, 827–830
- Infinity, 416
 - of infinities, 521
 - set theory as science of, 827–830
- Infinitesimal, the, 531
- Informal fallacies, 476
- Informal logic. *See* Logic, formal and informal
- Informal theories of interpretation, 410
- Informatics, 390–394
 - conceptual approaches to, 390–392
 - conceptual frameworks of, 392
 - etymology of, 390
 - theoretical constructions of, 392
 - theories of, 393
- Informational capital, 110
- Information exchange in patient-centered communication, 345
- Information-foraging theory, 393
- Information granularity, 458
- Information science, 843–844
- Information theory, 394–397
 - further developments in, 397
 - historical background, 397–404
 - Auguste Comte in, 403
 - Euclidean geometry in, 399
 - Francis Bacon in, 401
 - Galileo's defense of quantitative information, 400–401
 - Isaac Newton in, 402–403
 - from orality to literacy, 398–399
 - quantitative information and the modern world, 401–403
 - René Descartes in, 398, 400–402
 - scientific revolution as spirit of quantitative information in, 399–400
 - information and uncertainty in, 394–395
 - information between correlated systems, 395
 - operational treatment of information, 395–396
- Innovation hubs, 239
- Inscription devices, 837
- In Search of Lost Time*, 193
- Insight, intuitive, 412
- Instance-only models, 14
- Instinct, 412
- Institutional imperatives, 832
- Institutionalization phase of biostatistics, 107, 110–111
- Institutions of Medicine*, 564
- Instrumentalism, 404–406, 780
 - later developments in, 405
 - origin and significance of, 404
 - scientific realism and, 404–405
- Instrumental revolution, 127
- Integral calculus, 508
- Integral realism, 802
- Intended interpretation, 408
- Inter-coder reliability, 29
- Intercultural perspectives of cognitive science, 148–149
- Interests theory, 836
- Intergovernmental Panel on Climate Change (IPCC), 136
- Intermediate magma, 314
- Internal generalization, 306
- Internalism, 249–250, 425, 946–947
 - content, 662
 - rationality and, 776
- International Nucleotide Sequence Database Collection, 78
- International Organization for Standardization, 18
- International Vocabulary of Metrology*, 545
- Interpretation, 406–410
 - of empirical theories, 407–410
 - formal semantics of, 406–407
 - informal theories of, 410
 - of mathematical theories, 406–407
 - restricting, 410
- Interpretive flexibility, 837
- Interpretivism, 250–251
- Interquartile range (IQR), 866–867
- Intersubjectivity of measurement, 546
- Intra-coder reliability, 29
- Introduction to Logic*, 217, 363
- Introduction to Mathematical Studies and the True Reflections on the Four Unknowns*, 514
- Introduction to Meta-Mathematics*, 417
- Introduction to Metaphysics*, 417
- Introns, 78

- Intuition, 40, 411–415
 defined, 411
 facets of intuitive awareness and, 413–414
 factors affecting, 414
 features of intuiting and, 412–413
 interplay between sequential reasoning and, 414
 intuitive judgment and intuitive insight, 412
 as sensing plus sensemaking, 414–415
 what is not, 411–412
 Intuitionism in logic and mathematics, 415–420
 criticism of excluded middle in, 418–419
 finitism in, 416
 forms of, 416
 infinity and, 416
 intuitionistic logic and, 419
 ongoing development of, 419–420
 origins and history of, 417–418
 Invalid deductive arguments, 493–494
 Inverse-square law, 696–697
*Investigation of the Laws of Thought on Which are
 Founded the Mathematical Theories of Logic and
 Possibilities, An*, 292–293, 786
 Investing in infants, 264
 Irrationalism, 800
 Ishino, Yoshizumi, 73
 Islamic world. *See* Arab world
 Isomorphism, 113
 natural, 115
 unnatural, 115

 Jackson, Frank, 185
 Jacobi, Carl, 521, 708
 Jacopo, Mariano di, 235
 Jakobson, Roman, 151, 154, 160–161
 James, William, 404, 471–472, 660, 747–748, 928
 Janis, Irving, 777
 Janssen, Hans, 116
 Janssen, Zacharias, 116
 Jaśkowski, Stanisław, 630–631
 Jaspers, Karl, 952
 Jasso, Guillermina, 912
 Jayavarman VII, King, 569
 Jealousy, 264
 Jealousy protest, 265
 Jenner, Edward, 370, 582
 Jennings, Raymond, 630
 Jesus of Nazareth, 566
 Jevons, Williams, 219
 Jobs, Steve, 239
 Johannsen, Wilhelm, 81
 Johnson-Laird, Phillip, 584
 Johnstone, James, 942
 Joint method, 364
 Jordan, Pascual, 716

 Joseph, George Gheverghese, 323
 Joule, James Prescott, 709–710, 731
 Judgment
 engineering, 230
 intuitive, 412
 Julia, Gaston, 289, 539
 Jung, Carl Gustav, 434
 Justification (J), 423–428
 epistemic, 281–283
 foundationalism, coherentism, and, 425
 internalism, externalism, and, 425
 a priori J versus a posteriori J, 424–425
 theory of argumentation and, 426–428
 theory of J, 423–426
 warrant and, 946–947
 Justified true belief, 423, 433

 Kaccaya, Pakudha, 683
 Kador, Lothar, 91
 Kagan, Jerome, 194
 Kahneman, Daniel, 412, 476, 852
 Kaleb, Nicholas Nassim, 851–852
 Kan, Daniel M., 114
 Kanda, 683
 Kant, Immanuel, 8–9, 154, 249, 325, 664, 812, 898
 on concepts, 181
 influence of Newtonian theory on, 688
 non-Euclidean geometry and, 329
 phenomenalism and, 654
 rationality and, 775
 reasoning and, 784–785
 on reason setting goals and determining values, 471
 on temporal intuition, 415
 transcendental idealism, 32, 434
 on worldview, 950–951
 Kashi, Jamshid al, 529
 Kasparov, Garry, 41
 Kauffman, Stuart, 887
 Kavli prize, 607
 Keisler, Jerome, 641
 Kelland, Philip, 713
 Kelly, Benjamin, 835–836
 Kelvin, William Thomson, Lord, 88, 709–713
 Kempe, Alfred, 539
 Kendler, K. S., 189
 Kepler, Johannes, 4, 285, 400, 506, 530, 900–901
 abductive inference and, 788
 classical geometry and, 316, 320–321
 physical theory and, 676
 Renaissance physics and, 722, 724–726
 Kerlinger, Frank, 185–186, 189, 223
 Kiefer, Jack, 270–271
 Kierkegaard, Søren, 951
 Kim Il Sung, 46

- Kinematics, 769
- Kinetic (molecular) theory, 88, 429–432
 conceptual issues in, 430–432
 connecting mechanics and thermodynamics, 429–430
 need for discrete state spaces and, 431–432
 randomness and probability in ideal gas and, 431
- Kircher, Athanasius, 87
- Kirchhoff, Gustav, 126, 709
- Kleene, Stephen, 132, 290, 417, 420
- Klein, Felix, 328
- Klein, Greg, 411
- Kline, Rex B., 601
- Klug, Aaron, 93
- Knight, Frank, 851
- Knightian uncertainty, 851
- Knoll, M., 603
- Knorr-Cetina, Karin, 837
- Knowledge, 432–437
 abstract, 5–10
 acquiring, 434
 contemporary theories of medical, 555–556
 in context, 436–437
 defining, 432–433
 nature of, 433–434
 received, 771
 representing, 182, 435–436
 tripartite theory of, 249
 warrant and, 946
 of what to eat and what not to eat, 262
- Knowledge management, 436
- Knowledge worker, 436
- Kobayashi Maru game, 853
- Koch, Robert, 333–334, 575
- Koestler, Arthur, 190
- Koffka, Kurt, 583
- Kolander, Cheryl, 349
- König, Arthur, 712
- Kornberg, Hans, 69
- Koshland, Daniel, 66
- Kossel, Walther, 127
- Krebs, Hans Adolf, 69, 376–377
- Kresge, C. T., 604
- Krogh, August, 94
- Kronecker, Leopold, 416–417
- Kroto, Harry, 128, 604
- Kuhn, Arthur J., 177
- Kuhn, Thomas, 22, 45, 219, 280, 282–283, 286
 incommensurability thesis, 834–835
 on paradigm, 632–638
 on paradigm shift, 58–59
 reasoning and, 786
 revolutionary model, 669–670
 on scientific revolutions, 808–809
 on social nature of scientific change, 833–835
 theory change and, 898, 901, 903–904
 values in science and, 938–939
 on worldview, 953–954
- Kühne, Wilhelm, 64–67
- Külpe, Oswald, 284
- Kurtosis, 868
- Kuyper, Abraham, 951
- Kvanvig, Jonathan, 775
- Kyeser, Konrad, 235
- Laboratory ethnographies, 836–837
- Laboratory science, 19th century, 574–576
- Lactase, 69
- Laennec, Theophile-Hyacinthe, 573
- La Geometrie*, 507, 530
- Lagrange, Joseph-Louis, 501, 510, 512, 677, 693, 711–712, 768
- Lakatos, Imre, 221, 285, 634, 670
 on scientific progress, 794
- Lamarck, Jean-Baptiste, 80
- Lambert, Johann Heinrich, 326
- Lamé, Gabriel, 711
- Langford, C. H., 30
- Langley, Samuel Pierpont, 709
- Language
 abstraction and, 11–12
 abstract thinking and, 6–8
 causes of change in, 467–470
 cognitive science and, 145–146
 as constantly changing living organism, 467
 dual nature of linguistic oppositions and, 461
 functions of, 461
 grammatical change in, 469
 lexicon of, 821
 logic and, 485–486
 mathematics, cybernetics, and, 203–204
 natural, 474
 proto-, 466
 sound change in, 468–469
 vocabulary change in, 469–470
- Laozi, 436, 773
- Laplace, Pierre-Simon, 288–290, 501, 506, 510–512, 693, 754, 851
- Laplace-Gauss distribution, 290
- Laplace's demon, 754
- Last universalist, 523
- Latin Squares, 540
- Latour, Bruno, 22–23, 837–838
 on nonhuman actants, 838
- Laudan, Larry, 634, 670–671, 759–760, 762
- LaVeist, Thomas, 347–348
- Lavoisier, Antoine, 125, 511, 693
- Law, abduction in, 3
- Law of conservation of matter, 125

- Law of contradiction, 498
- Law of gravity, 98
- Law of inertia, 724
- Law of large numbers (LLN), 854, 868–869
- Law of nature, 440
- Law of propagation of uncertainty, 20
- Law of requisite variety, 887
- Law of Thought, The*, 219
- Laws, scientific, 439–446
 - combination views of, 445
 - dispositional essentialism, 443–445
 - new directions in, 445
 - nomoc necessitation, 441–443
 - pragmatism and, 445
 - simple regularity view, 439–440
 - sophisticated regularity view, 440–441
 - structuralism and, 445
- Laws of nature, 446–450, 916–917
 - best-systems accounts, 447–448
 - causal/dispositional accounts, 448
 - constraint accounts, 448–449
 - counterfactual accounts, 448
 - nomological realism, 446–447
 - as primitive, 447
 - regularity accounts, 447–448
 - semantic conception of theories and, 916–917
 - statistical, 449–450
 - universal accounts, 446–447
- Lawvere, Bill, 114
- Laymon, Ronald, 916
- Lebedev, Pyotr Nilolayevitch, 709
- Le Bel, Joseph Achille, 126
- Lebesgue, Henri, 417
- LeBlanc, Nicolas, 125
- Lederberg, Joshua, 37
- Le Dran, Henri-Francois, 562
- Lee, H., 604
- Legal-rational system of authority, 46
- Legendre, Adrien Marie, 268, 501, 693
- Legionella*, 371
- Lehrer, Jonah, 193
- Lehrman, Daniel S., 225
- Leibniz, Gottfried Wilhelm, 173, 219, 702, 712, 768, 849
 - Boolean algebra and logic and, 504
 - panpsychism and, 660
 - rationality and, 775
 - on tangency, maxima, minima, and estimates, 531
 - theory of calculus, 508
- Lemaître, Georges, 197–198
- Leonelli, Sabina, 210–211
- Leoniceno, Niccolò, 578
- Leplin, Jarrett, 802
- Leptokurtic curve, 868
- Leray, Jean, 537
- Lerner, Richard M., 225
- Leucippus, 682
- Leung, Kwok, 193
- Level of analysis, 913
- Leviathan*, 550
- Levin, Leonid, 291
- Lewes, George Henry, 174
- Lewin, Kurt, 583–584
- Lewis, C. I., 497, 630–631, 661
- Lewis, David, 589
- Lewis, Gilbert N., 127
- Lewontin, Richard, 618–619, 942–943
- Lexemes, 821
- Lexical semantics, 818
- Lexicon of a language, 821
- Lexis, 457
- Liar paradox, 512, 640, 643
- Lie, Marius Sophus, 508
- Liebig, Justus von, 63–65, 126, 574, 940
- Life history theory, 261
- Life sciences, contemporary, 450–456
 - computer simulations in, 453–454
 - data, dynamics, and computers in, 452–455
 - history, philosophy, and theory of, 451–452
 - nature, science, and scientists in, 450–451
 - scientific animation in, 454–455
- Life-spheres, 776
- Lightning attractor, 696
- Limestone, 314
- Limitation theorems, 335
- Linacre, Thomas, 578
- Lindley, Dennis, 858
- Linear models, 597–598
- Linear systems, 887
- Linguistic analysis, levels of, 461–462
- Linguistic annotation, 458
- Linguistic appropriations of worldview, 954–955
- Linguistic determinism, 461
- Linguistic frameworks, 456–460
 - applications and usage of, 457–458
 - building and implementation of, 456–457
 - fields that apply, 458–459
 - information granularity in applying, 458
 - instances of, 457
 - used in linguistic annotation, 458
- Linguistic perspectives on cognitive science, 145–146
- Linguistic rationalism, 385–386
- Linguistic relativism, 460–461
- Linguistics
 - contemporary, 460–464
 - dual nature of language and linguistic oppositions in, 461
 - functions of language and, 461
 - hybrid disciplines in, 463–464

- language features in, 460
- levels of linguistic analysis in, 461–462
- linguistic relativism versus language universal in, 460–461
- historical, 465–470
 - goals and methods of, 465–466
 - language as constantly changing living organism and, 467
 - language change and its causes, 467–470
 - reconstructing a proto-language in, 466
 - relations to other disciplines, 470
 - typological reconstruction in, 466–467
- Linnaeus, Carl, 79–81, 563, 893–895
- Lipids, 68, 73
- Lipmann, Fritz Albert, 69
- Lippmann, Gabriel, 714
- Lister, Joseph, 332–333, 573, 713
- Lithification, 314
- Lithosphere, 310–311, 741–742
- Little, Arthur D., 127
- Little Balance, The*, 723
- Littlejohn, Stephen, 948
- Li-Tzu Hu, 601
- Liu Hui, 505
- Living nature of language, 467
- Living systems, 888
 - physics of, 98–99
- Lloyd, Elizabeth, 916
- Lobachevskii, Nikolai, 325–328, 519, 523–524, 629, 782
 - on hyperbolic geometry, 796
- Locke, John, 110, 325, 358, 434, 436, 812
 - bundle theory and, 589
 - on capitalism, 610–611
 - capitalism and, 243
 - on formation of concepts, 181
 - rationality and, 774
 - reasoning and, 786
 - religion and, 698–699
 - thought experiments, 924–925
- Locution, 462
- Locutionary act, 817
- Logarithm, 530–531
- Logic
 - Aristotle's, 470–471, 784
 - axiom schemas in, 48–50
 - Boolean, 219, 292–293, 786
 - categorical, 496–497
 - completeness and, 169–170
 - in contemporary times, 484–485
 - deductive reasoning and inductive reasoning and, 292–293
 - five branches of, 496–498
 - formal and informal, 293, 470–477
 - deductive, 473–477
 - one hundred definitions in, 471–473
 - soundness in, 848–849
 - foundations of, 291–293
 - fuzzy, 178–179
 - of generalization, 305
 - of hypothesis testing, 350–351
 - inductive, 477–483, 899 (*See also* Induction)
 - critique of, 479–481
 - justification of, 481–482
 - origins and development of, 478–479
 - in the social sciences, 482
 - inferentialism applications in, 386–388
 - intuitionism in mathematics and, 415–420
 - intuitionist, 419
 - mathematical, 288, 470, 492
 - modal, 497–498
 - necessity in, 292
 - of nonsense, 630
 - paraconsistent (*See* Paraconsistency)
 - of paradox, 631
 - predicate, 292, 497
 - reasoning and, 784
 - rival conceptions of, 293
 - soundness in, 846–847
 - symbolic, 492, 497
 - truth-functional, 497
 - value of, 498
- Logical conjunction, 472–473
- Logical connectors, 485
- Logical empiricism, 900–901
- Logical form, 847–848
- Logical nativism, 485
- Logical operators, 485
- Logical positivism, 590, 801, 817
- Logical reasoning, 2
- Logical Structure of Mathematical Physics, The*, 919
- Logical theory, 492–499
 - basics of, 493
 - birth of, 492
 - categorical logic in, 496–497
 - consistency, implication, and equivalence in, 495
 - deductive and inductive arguments in, 493–495
 - five branches of logic and, 496–498
 - informal logic in, 498
 - modal logic in, 497–498
 - necessity and contingency in, 495–496
 - predicate logic in, 497
 - truth-functional logic in, 497
 - value of logic and, 498
- Logic and language, 483–491
 - Aristotle's logic or traditional logic, 483–484
 - atypical population, 490–491
 - in contemporary times, 484–485
 - details of good testing in, 474–475

- evidence from cross-cultural child and adult language research, 486–488
- Modularity Matching Model (MMM), 488–489
- negation, disjunction, and quantificational words, 489–490
- relationship between, 472–473
- Logicism, 33, 336
- Logic of Scientific Discovery, The*, 284–285
- Logics of formal inconsistency (LFIs), 631
- Logic Theorist*, 36
- Logic Theory Machine*, 36
- Logik der Forschung*, 284
- Logos, 773
- Logotherapy, 59
- Longino, Helen, 761
- Longitudinal studies, 256
- Lorentz, Hendrik, 736
- Lorentz transformations, 688
- Lorenz, Edward Norton, 176, 539
- Lorenz, Konrad, 225
- Loschmidt, Johann Joseph, 735
- Lotka, Alfred, 34
- Lotka-Volterra model, 34
- Louis, Pierre-Charles-Alexandre, 573
- Lovelock, James, 128, 245, 943
- Lowe, Jonathan, 590
- Lowspread, 866
- Ludwig, Karl Friedrich Wilhelm, 574
- Luhmann, Niklas, 176, 889
- Łukasiewicz, Jan, 292, 629–630
- Luminiferous aether, 736
- Lusk, Greg, 211
- Lynch, Michael, 837–838
- Lyotard, Jean-François, 672
- Lysosomes, 121
- Lysozyme, 69
- Mach, Ernst, 405, 771
- phenomenalism and, 655
- Machine state functionalism, 658
- Mackie, John Leslie, 555, 590
- Mac Lane, Saunders, 113–115
- Maclaurin, Colin, 693, 712
- Macroevolution, 83
- Maestlin, Michael, 725
- Mafic magma, 314
- Magical thinking, 28
- Magma, 314
- Magnetic resonance imaging (MRI), 626–627
- Magnetism, 726
- Magnus, H. G., 709
- Male-centered patriarchal worldview, 777
- Malthus, Thomas, 612–613
- Malus, Étienne-Louis, 713
- Management science and software engineering, 844
- Manhattan Project, 536
- Mann, Horace, 436
- Mann, James E., 644
- Mannheim, Karl, 836
- Mantle, Earth's, 310
- Mapping account of applied mathematics, 33–34
- Map-territory problem, 107
- March, James, 777
- Marconi, Guglielmo, 236
- Marcos, João, 631
- Marcuse, Herbert, 814
- Mareschal, Charles Georges, 562
- Margulis, Lynn, 83, 245, 614
- Mariotte, Edme, 711
- Market failure, 233
- Markov, A. A., 420
- Markov strategy, 300
- Martin, Emily, 838
- Martini, Francesco di Giorgio, 235
- Marx, Karl, 220, 583, 688, 814, 836, 951
- Marxism, 198, 284, 611, 688, 791, 814–815, 951–952
- Material Basis of Evolution, The*, 83
- Material inference, 376–377
- Materialism, 774
- Material relationality, 23
- Materials science, 727
- Mathematical Analysis of Logic, The*, 219
- Mathematical and Philosophical Dictionary*, 545
- Mathematical constructivism, 782
- Mathematical linguistics, 459, 463
- Mathematical logic, 288, 470, 492
- soundness and completeness in, 849
- Mathematical Principles of Natural Philosophy. See Principia Mathematica*
- Mathematical proof, 782
- Mathematics
- 19th century, 517–525
- Bernhard Riemann and, 519–520
- Boolean algebra and logic in, 520–521
- Carl Friedrich Gauss and, 518–519
- Évariste Galois and, 517
- Georg Cantor and, 521–522
- Henri Poincaré and, 522–523
- other notable mathematicians of, 523–524
- 20th century, 532–541
- in the 1970s and beyond, 539–541
- Alan Turing and, 535–536
- Alexander Grothendieck and, 537
- Andre Weil and, 536–537
- Bertrand Russell and Alfred North Whitehead and, 534
- David Hilbert and, 533–534
- G. H. Hardy and Srinivasa Ramanujan and, 532–533
- John von Neumann and, 536

- Julia Robinson and Yuri Matiyasevich and, 538–539
- Kurt Gödel and, 535
- Paul Erdős and, 537–538
- Paul Joseph Cohen and, 538
- abduction in, 3
- abstraction in, 13–14
- accuracy of computation in, 17–18
- algebraic theory in, 288
- algorithms and data structures, 290
- antiquity, 501–506
 - categorization by geography, 518–506
 - importance of historical artifacts for understanding, 501–518
 - practical and theoretical mathematics in, 518
- applicability in science, 31–35
- axiom schemas in, 50–51
- as basis of the physical sciences, 288
- beginning of mathematical sciences and, 289
- beginnings of artificial intelligence and, 289
- emergence of neuroscience and, 290–291
- foundations of modern-day, 288–291
- game theory and, 289
- implications of non-Euclidean geometry for, 329–331
- intuitionism in logic and, 415–420
- language, cybernetics, and, 203–204
- mathematical logic in, 288
- number theory and, 289
- of perturbation theory, 645–650
- realism in, 784–785
- set theory as foundation of, 830–831
- software engineering and, 843
- speculation in, 851–852
- statistics and probability, 289–290
- structure of quantum theory, 716–717
- Mathematics, Enlightenment, 506–513
 - development of analysis, number theory, and mechanics in, 510
 - early, 506–507
 - early 18th-century, 509–510
 - foundational, 507–508
 - greenhouse effect and, 511–512
 - later Renaissance mathematics and theoretical explanation of the Solar System in, 510–511
 - legacy of, 512
 - theory of calculus in, 508–509
- Mathematics, Middle Ages, 513–517
 - significant Arabic mathematicians of, 513–514
 - significant Chinese mathematicians of, 514
 - significant European mathematicians of, 514–516
 - significant Indian mathematicians of, 516
- Mathematics, Renaissance, 525–532
 - arithmetic and algebraic studies in, 527–529
 - astronomy and trigonometry in, 529
 - geometry in, 529–530
 - Greek mathematics and, 525–526
 - the infinitesimal and the theory of fluxions and, 531
 - influence of Islamic tradition on, 526–527
 - logarithm and, 530–531
 - number theory and, 530
 - scientific renaissance and, 527
 - Scientific Revolution and, 529–530
- Mathematics: From the Birth of Numbers*, 827
- Matiyasevich, Yuri, 522, 538–539
- Matrices, 190
- Matter theory, 541
- Matural abstraction, 12–13, 81
- Maudslay, Henry, 236
- Maupertuis, Pierre Louis, 693
- Maxmin strategy, 299
- Maxwell, James Clerk, 88–90, 688, 700, 709, 716, 731, 734–736, 938
- Mayr, Ernst, 82
- McCarthy, John, 36, 177
- McClintock, Barbara, 83–84
- McCombs, Maxwell, 187
- McCulloch, Warren, 38, 177, 201–203, 207
- Mead, Margaret, 152, 154
- Meade, Richard, 76
- Mean, 865–866
- Meaning, personal, 933–934
- Meaning of living form, 207
- Meanings as concepts, 818
- Mean velocity theorem, 707
- Measurability, 544–546
- Measurand, 19
- Measurement, 541–547
 - accuracy of, 18–21
 - background on, 542–544
 - central tendency, 864–866
 - dispersion, 866–867
 - indirect, 546
 - measurability and, 544–545
 - more encompassing concept of measurability and, 545
 - of nonphysical properties, 545
 - of nonquantitative properties, 544–545
 - as representation, 543–544
 - scope of theory of, 541–542
 - theory of process of, 545–546
 - uncertainty in, 546
- Mecanique Analytique*, 510, 711
- Mecanique Physique*, 712
- Mechanical engineering, 547
- Mechanical worldview, 321–322
- Mechanicism, 770–771
- Mechanics, 510
- Mechanics, Or the Science of Motion, Analytically Expounded*, 768
- Méchanique Analytique*, 693

- Mechanism, 592
- Medical biophysics, 96
- Medicalization, 111, 553
- Medical knowledge and evidence, theories of, 555–556
- Medical sciences, abduction in, 2–3
- Medicine
- 19th century, 569–577
 - age of systems and, 547–549
 - the laboratory in, 574–576
 - presenting alternatives in, 576
 - role of the clinic in, 571–574
 - 20th century (*See* Medicine, 19th century; Medicine, contemporary)
 - antiquity, 547–552
 - in China, 550–551
 - in Egypt, 547–548
 - in Greece and Rome, 548–550
 - in India, 551
 - in Mesopotamia, 547
 - contemporary, 552–558
 - reductionism and biomedical model in, 556–557
 - theories of causation in, 554–555
 - theories of health and diseases in, 552–554
 - theories of medical knowledge and evidence in, 555–556
- Medicine, Enlightenment, 558–565
- anatomy in, 560–561
 - humoral theory of, 558–560
 - nosology and disease theory in, 563–565
 - surgery in, 561–563
 - theories of health and body in, 558–561
- Medicine, Middle Ages, 566–569
- in the Arab world, 567
 - beginnings and early practices in European, 566–567
 - in China and Southeast Asia, 568–569
 - the plague and its aftermath and, 332, 369, 568
- Medicine, Renaissance, 577–583
- beginnings of, 578
 - contagious diseases, poisons, and remedies in, 579–580
 - further outbreaks of the plague and, 581
 - outside Europe, 581–582
 - universities, medical research, and religious reform and, 578–579
 - war injuries and trauma in, 580
 - William Harvey and discovery of circulation of blood and, 580–581
- Meditations on First Philosophy*, 249, 401–402
- Mellers, Barbara, 854
- Melloni, Macedonio, 709
- Memory, 435
- Mendel, Gregor, 81–82, 110, 261, 746, 763, 942
- Mendeleev, Dmitri I., 126
- Mendelian genetics, 81–82
- Mendenhall, Thomas C., 712
- Meno*, 778–779
- Mental models, 277, 583–586
 - in communication and teaching, 585–586
 - executive functions, worldviews, and statistics in, 584–585
 - origins and development of, 583–584
 - principles of, 584–585
- Mental spaces, 190–191
- Mental state eliminativism, 657–658
- Menten, Maud, 66
- Merleau-Ponty, Maurice, 952
- Mersenne, Marin, 531
- Mersenne prime numbers, 540
- Merton, Robert, 755, 761, 832, 852, 936–937
- Merton College, 707
- Mesokurtic curve, 868
- Mesopotamia
 - mathematics in, 503–504
 - medicine in, 547
- Mesosphere, 311
- Messenger RNA (mRNA) vaccines, 73–74
- Meta-analysis, 369
- Metabolism, 69–70, 75
- Metacognitive knowledge, 436
- Metamorphic rocks, 314–315
- Metaphors, 280
- Metaphysical glue, 447
- Metaphysical naturalism, 590
- Metaphysical posits, 445
- Metaphysics, 587–591, 917
 - background on, 587–588
 - bundle theory and, 589
 - fall and rise of, 589–590
 - influence of, 590–591
 - of mental phenomena, 656–661
 - objects and properties in, 588–589
- Metaphysics*, 32, 544, 583, 927
- Metatheory, 591–594
 - behind theory and method, 592–593
 - evaluating theory using, 593–594
 - mapping the conceptual sense of, 591–592
- Meteorology. *See* Climate science
- Methicillin-resistant *Staphylococcus aureus* (MRSA), 371
- Method and methodology, 153–163
 - metatheory behind, 592–593
 - philosophy of science and, 663–665
- Method of agreement, 364
- Method of authority, 248
- Method of Curing Fevers, The*, 563
- Method of difference, 364
- Method of science, 248
- Method of tenacity, 248
- Method of Treatment for Wounds Caused by Firearms, The*, 580

- Methodological theses in scientific realism (SR), 807
 Methodology of scientific research programs (MSRPs), 221
 Michaelis, Leonor, 66
 Michaelis-Menten kinetics, 92
 Michel, Hartmut, 94
 Michell, John, 693
 Michelson, Albert A., 714, 736
 Microbe hunters, 575
 Microevolution, 83
 Microzymas, 333
 Middle Ages, the
 mathematics of, 513–517
 medicine in, 566–569
 physics of, 703–707
 reasoning in, 784–785
 Mid-spread, 866
 Milgram, Stanley, 47–48, 777
 Military engineering, 235–236
 Mill, John Stuart, 32, 221, 363–365, 611
 canons of induction, 479
 phenomenalism and, 654–655
 rationality and, 774
 reasoning and, 786
 Miller, James G., 885
 Miller, William Hallows, 89
 Millikan, Robert, 781
 Mills, Robert, 686
 Mill's rules, 498
 Mind-brain identity theory, 771
 Mind-dependent laws, 441
 Mind-(in) dependence, 445
 Mind/matter dualism, 105
 Minerals and rocks, 313–315
 Minimal groups paradigm, 266
 Minimalism, 928–929
 Minimal mechanistic models, 99
 Minkowski, Hermann, 524
 Minkowski space-time, 524
 Minority health, 347–348
 Minsky, Marvin, 38–40
Mirifici Logarithmorum Canonis Descriptio (A Description of the Marvelous Rule of Logarithms), 530–531
 Misak, Cheryl, 928
 Mishler, Elliot, 344
 Mitigation climate policy, 141–142
 Mitochondria, 76, 121
 Mitroff, Ian, 832
 Mitscherlich, Eilhard, 65
 Mixed strategy, 299
 Mobile resources, 609
 Möbius, August F., 524, 712–713
 Modality, 445
 Modal logic, 497–498
 Mode, 866
 Model building, 583–586
 Modeling, 594–602
 categorical, 597
 how to conduct, 595–596
 linear, 597–598
 nonlinear, 599
 semantic rules of, 817, 914–915
 structural equation (SEM), 599–601
 when to use, 595
 why conduct, 596–597
 Model systems, 99
 Model-theoretic argument, 408
Model Theory, 641
 Modern-day logic, 497
 Modern synthesis, 261, 763
 ModSyn, 82–83
 Modularity Matching Model (MMM), 488–489
 Modus ponens, 2
 Moerner, W. E., 91
 Moigno, François, 711
 Molecular biophysics, 92–94
 Molecular genetics, 70–71
 Molecular modeling, 91
 Molecular neuroscience, 623
 Molina, Mario J., 128
 Momentum, 769
 Monomers, 68
 Monophthongization, 468
 Monstrous moonshine, 540
 Monte Carlo simulations, 453
 Moore, Don, 854
 Moore, G. E., 30, 590
 Moore, Omar Khayyam, 301
 Moral rationalism, 774
 Moral sense and abstract knowledge, 8–9
 Morgagni, Giovanni Battista, 559
 Morgan, Thomas Hunt, 81
 Morgenstern, Oskar, 299, 301, 853
 Morley, Edward W., 736
 Morphisms, 114–115
 Morphology, 462
 Morris, Charles, 637
 Morton's fork paradox, 850
 M theory, 675, 678, 692
 Muir, John, 244
 Müller, Erwin, 604
 Müller, Johannes Peter, 527, 565
 Müller, K. Alex, 730
 Mulliken, Robert S., 128
 Mullis, Kary, 70
 Multifactorial accounts, 760–761
 Multimodal distribution, 866
 Multiple Intelligences, 435, 586, 932–933
 Mumby, Dennis, 777

- Murphy, Robert, 712
 Musk, Elon, 239
 Muslim world. *See* Arab world
- Naess, Arne, 248
 Naïve realism, 280, 799, 802–803
Naïve Set Theory, 827
 Nanochemistry, 128
 Nanocrystalline, 604
 Nanoparticles, 603
 Nanorobotics, 606
 Nanoscience, 605
 Nanotechnology, 603–608
 applications of, 606–607
 as expanding field, 605
 Kavli prize in, 607
 key developments in, 604–605
 layers and materials in, 605–606
 origins of, 603–604
 in science, economics, and politics, 607–608
 Nanotubes, 606
 Napier, John, 506, 530–531
 Nash, John, 299, 301, 540, 853
 Nash equilibrium, 299–300, 540, 853
 National Institute of Genetics, Mishima, Japan, 78
 Naturalism, 12–13, 553
 explanation and, 277
 metaphysical, 590
 Naturalistic fallacy, 262
 Naturalists, 553, 617, 672
 Natural language, 474
 Natural philosophy, 680
 Natural resources, 609–616
 capitalism and accumulation of, 610–611
 common pool resources and, 614
 dominance and, 610
 Gaia theory and, 614–615
 Malthus, overpopulation, and neo-Malthusians and, 612–613
 resource law and, 615
 scarcity, opportunity cost, and externalities of, 611–612
 Tragedy of the Commons and, 613–614
 types of, 609–610
 Natural selection, 80–82, 260–261, 451, 616–622
 as cause of evolution, 619–620
 Darwinism in historical context and, 620–621
 as kind of evolution, 617–619
 principle of, 616–617
Nature of Necessity, The, 498
 Naugle, David, 951–952
 Navier, Claude-Louis, 512, 711
 Navier, H., 768
 Neale, M. C., 189
 Necessary falsehood, 495–496
 Necessary truth, 495–496
 Necessity in logical theory, 495–496
 Negation, 489–490
 syntax of, 883–884
 Negative entropy, 888
 Negative feedback, 888
Nei Ching, 551
 Neo-Darwinism, 261, 764
 Neologicism, 33
 Neo-Malthusians, 612–613
 Neovitalism, 941
 Nerve impulse, 92
 Neuberg, Carl, 68
 Neumann, C., 712
 Neumann, Franz Ernst, 709
 Neural excitation, 181
 Neural networks, 453
 Neurath, Otto, 634
 Neuro-aesthetics, 624
 Neurochemistry, 71
 Neuroeducation, 624
 Neuroethics, 624
 Neurolinguistics, 458, 463, 624
 cognitive science and, 146
 Neuromanagement, 624
 Neurophysiology, 87
 Neuroscience, 290–291, 622–628
 approaches and tools used in, 626–627
 areas of research in, 623
 background on, 622
 cognitive science and, 149
 future of, 627
 and other disciplines, 623–624
 research in, 624–626
Neurospora crassa, 82
 Neutral monism, 660
 Newcomen, Thomas, 236
 Newell, Allen, 36
New Organon, 401
 New riddle of induction, 480
 Newton, Isaac, 32, 87, 98, 110, 194, 283, 288, 398–399, 682, 713, 812
 classical mechanics, fundamental forces and, 676–677
 cosmology, 195
 foundational mathematics, 507–508
 geometry and, 316
 on heavenly movements, 321–322, 325, 699
 in historical background of information theory, 402–403
 on the infinitesimal, 531
 laws of motion, 447–449
 rational mechanics and, 767–768
 religious values of, 937
 Renaissance mathematics and, 529
 as theoretical physicist, 675

- theory change and, 900–901
 three-body problem and, 522–523
 Next Generation Science Standards, 436
 Neyman, Jerzy, 268, 350
 Neyman-Pearson lemma, 269
 Neyman-Pearson test, 873
 Neyman structure, 873
 Nicholas, Ernest Fox, 709
 Nicholson, Daniel J., 592
 Nietzsche, Friedrich, 196, 775–776, 929, 949
 Nihilism, 12
Nine chapters, 504–505
 Nineteenth century, the
 biochemistry in, 63–68
 biophysics in, 86–90
 mathematics in, 517–525
 medicine in, 569–577
 physics in, 707–715
 reasoning in, 786
 NMR spectroscopy, 91, 93
 Noether, Emmeline, 538
 Noise, 211–212
 Nollet, Abbé J. A., 87
 Nomic necessitation view of scientific laws,
 441–443
 Nominalism, 12, 588, 780
 Nomological realism, 446–447
 Nondeterministic polynomial time, 291
 Nonhuman actants in social construction of scientific
 knowledge, 838
 Nonjustificationism, 423–424
 Nonlinear and chaotic systems theory, 740
 Nonlinearity in complex systems, 174–175
 Nonlinear models, 599
 Nonlinear systems, 887
 Nonparametric test statistics, 351–3430
 Nonrational models, 852
 Nonrenewable resources, 609
 Nonresistant measure, 866
 Nonresistant statistic, 866
 Nonscience, 663, 665, 762
 Nonsense logic, 630
 Nonstatement view, 912
 Nordström, Gunnar, 699
 Normal distribution, 867
 Normal science, 635
 Normativism, 553
 Northrop, John Howard, 69
 Norton, John, 921
 Norvig, Peter, 39
 Nosology, 563–565
 Nostratic language, 466
Notes on Constructive Mathematics, 420
 Novel, Alfred, 126
Novum Organum Scientiarum, 364, 479, 785
 Nuclear envelope, 121
 Nuclear magnetic resonance (NMR), 90
 Nuclear pores, 122
 Nucleic acids, 73–74
 Nucleoplasm, 120, 122
 Nucleus, cell, 121–122
 Null hypothesis significance testing (NHST), 377
 Null hypothesis testing, 357, 855, 869–870
 Number theory. *See* Mathematics, 19th century;
 Mathematics, 20th century
 Numerical indicators, 864
 Nunnally, Jum, 186

 Obedience to authority, 47–48, 777
*Obedience to Authority: An Experimental
 View*, 47
 Objectivism, 251
 Objectivity, 386, 480
 of measurement, 546
 Object related measurement, 546
 Objects and properties in metaphysics, 588–589
 O'Brien, S., 604
 Observation, 663
 Observational unit, 864
 Observed values, 864
 Observer effect, 852
 Oceanic crust, 313
 Ockham's razor, 12, 97, 589, 674, 707, 850
 Ockham's rug, 99
Odyssey, 398
 Ogden, Charles K., 181
 O'Hear, Anthony, 478
 Ohm, Georg, 708
 Olsen, Wendy, 803
 Olszewski, Karol, 710
 Olthuis, J. H., 956
On Architecture, 319
 One-tail test, 352–354
On Floating Bodies, 683
 Ong, Walter J., 398–399, 951
On Interpretation, 470
On the Calculation with Hindu Numerals, 513
On the Origin of Species, 80–82, 110, 260, 616–617,
 787, 923
On the Revolutions of the Heavenly Sphere, 320,
 399, 721
On the Structure of the Human Body, 578
On the Theory of Scales of Measurement, 543
On the Use of the Parts, 560
 Ontic depth, 803
 Ontic structural realists, 445
 Ontological realism, 805
 Ontological theses in scientific realism (SR), 804–805

- Ontology, 249, 323, 435
 of data and data models, 213–214
 process, 592
 substance, 592
On Triangles of Every Kind, 515–516, 527
 Open data movement, 216
Opening of the Universe, The, 516
 Open systems, 173, 886
 Operations research, 295–296
 Operator error, 232–233
 Oppenheim, Paul, 751
 Opportunity cost, 612
 Optics, 88, 713–714
 Optimum experimental design, 270–271
Opus Palatinum de Triangulis, 529
Orality and Literacy, 398
 Order of explanation, 444
 Ordinality, 780, 826
 Ordinary language philosophy, 276, 590
 Oresme, Nicole, 705, 707
 Organelles, 117
 Organicism, 592, 941–942
 Organismic biology, 885
Organization of Behavior, 290
 Organized skepticism, 761, 936
Organon, The, 363, 470, 534
Origin of Continents and Oceans, The, 743–744
 Orr, James, 951
 Orsted, Hans Christian, 708
 Ortelius, Abraham, 745
 O'Shaughnessy, John, 222–223, 368
 Osmotic pressure, 87
 Ostensive interpretation, 408
 Osteopathy, 576
 Ostrom, Elinor, 614
 Ostwald, Friedrich Wilhelm, 127, 770
Our Bodies, Ourselves, 349
 Outer core, 311
 Outliers, 866
 Overinterpretation, 28
 Overpopulation, 612–613
 Overreading, 28
 Overton, William, 285, 592
 Oyama, Susan, 226

 Pace, Norman, 73
 Pacioli, Luca, 528, 850
 Paganini, Nicolo, 524
 Palatalization, 468
 Paleontology, 310
 Palitzsch, Johan Geor, 695
 Panexperientialism, 659–660
 Pantheism, 774, 785

 Paraconsistency, 629–632
 adaptive logics and, 631
 beginnings of, 629–630
 dialetheism and, 631
 logics of formal inconsistency (LFIs) and, 631
 main approaches to, 630–632
 paraconsistent logic, 498, 629–630
 preservationism and, 630–631
 relevant logics and, 631
 Paraconsistent logic, 498, 629–630
 Paradigm, 45, 632–638, 669, 833–834, 902–903, 905
 development of science and, 635–636
 by disciplinary matrix, 634
 further developments in, 637–638
 as interpretive constraints, 834
 introduction to Kuhn's, 633
 Kuhnian conception in the context of 20th-century philosophy of science, 633–634
 Kuhn's concept(s) of, 634–635
 Kuhn's relations with classical philosophy of science and, 636–637
 shift of, 58–59
 Paradoxes, 534, 638–644
 of analysis, 30
 examples of, 639–641
 importance of, 642–643
 liar, 512, 640, 643
 logic of, 631
 problems and solutions with, 641–642
 Russell's, 219–220, 640–642, 830
 Sorites, 639–640, 642
 Zeno's, 639, 642
 Paralinguistic features, 463
 Parallelists, 661
 Parameters, 867
 Parametric rationality, 301
 Parametric studies, 867
 Parametric test statistics, 351–3430
 Paré, Ambroise, 580
 Parental investment theory, 263
 Parenting and evolutionary psychology, 264–265
 Parent language, 468
 Pareto optimality, 101
 Pargeter, Edith, 567
 Parmenides, 12
 Parsing, 879
 Parsons, Talcott, 176, 889
 Partial orbs, 705
 Particle physics, 536, 737–738
 Pascal, Blaise, 530, 772, 785, 850–851
 Pask, Gordon, 177, 207
 Pasqualin, Renata, 605
 Pasteur, Louis, 23, 63–65, 126, 332, 334, 370, 575
Pasteurization of France, The, 22

- Paternity uncertainty, 264
 Path analysis, 601
Pathwaie to Knowledge, The, 528
 Patient-centered communication, 344–347
 Patient self-management, 346
 Pattern generators, 15
 Pauli, Wolfgang, 405, 689, 737
 Pauling, Linus, 92, 128
 Payen, Anselme, 64
 Peacock, George, 524
 Peaked curve, 868
 Peano, Giuseppe, 524, 816, 849
 Peano arithmetic, 171
 Pearce, Jone L., 411, 413
 Pearson, Karl, 110, 350, 786
 Pedagogy of the oppressed, 436
 Peebles, James, 199
 Peirce, Charles Sanders, 1–2, 153, 219, 295, 380, 747, 928
 on abductive reasoning, 788
 Boolean logic and, 520
 hundred definitions of logic and, 471–472
*Pendulum Clock: Or Geometrical Demonstrations
 Concerning the Motion of Pendula*, 768
 Pepper, Stephen C., 592
 Perceptual entitlement, 947
 Perelman, Grigori, 540–541
 Perfect cosmological principle, 198
 Perkin, William Henry, 126
 Perlocution, 462
 Peroxisomes, 121
 Perpetual motion machine of the first kind, 734
 Perry, William G., 934
 Personalized medicine, 557
Personal Knowledge, 953
 Personal understanding, 933–934
 Perspectivalism, 775–776
 Perturbation theory, 644–652
 limitations of, 652
 mathematical background of, 645–650
 regular perturbations, 644
 singular perturbations, 650–652
 use of series in, 648–650
 Petit, Alexis-Thérèse, 709
 Petit, Jean-Louis, 562
 Petrology, 310
 Pfeffer, W. F. P., 87, 710
 Phenomenalism, 279–281, 653–656, 800
 Bertrand Russell and, 655
 David Hume and, 654
 Ernst Mach and, 655
 George Berkeley and, 653–654
 Immanuel Kant and, 654
 John Stuart Mill and, 654–655
 Phenomenology, 250–251
 Phenotype, 81
 Phillips, David Chilton, 69
 Philoponus, John, 705
 Philosopher's Football, 540
Philosophia Botanica, 79
Philosophiae Naturalis Principia Mathematica (Principia),
 32, 321–322, 399, 507, 529, 688, 767
Philosophical Essay on Probability, A, 511
Philosophical Foundations of Physics, 52
 Philosophical linguistics, 459, 464
 Philosophical naturalism, 775
 Philosophical perspectives of cognitive science, 148
Philosophie Zoologique, 80
 Philosophy, concepts in, 181, 526
 Philosophy of mind, 148, 656–663
 approaches to metaphysics of mental phenomena in,
 656–661
 brain state-mental state identity theory in, 657
 central issues in metaphysics of mental phenomena and,
 661–662
 content internalism and externalism in, 662
 dualism in, 660–661
 embodied cognition in, 658–659
 extended mind theory in, 659
 functionalism in, 658
 mental state eliminativism in, 657–658
 neutral monism in, 660
 panexperientialism in, 659–660
 qualia and, 661–662
 Philosophy of science, 663–673
 abduction in, 3
 demarcation problem of telling science from nonscience
 and pseudoscience in, 663
 induction and deduction in, 665–668
 Kuhnian conception of paradigm in context of, 633–634
 Kuhn's relations with classical, 636–637
 natural and social sciences, 671–672
 observation in, 663
 principles of scientific reasoning in, 663–665
 scientific progress, revolution, evolution, programs, and
 traditions in, 668–671
 scientific realism (SR) and, 800–802, 804
 scientific revolutions and, 812–813
 theories and methods in, 663–665
 value-free science and scientific authority, 672–673
Philosophy of Set Theory, The, 416
Philosophy of the Inductive Sciences, The, 192
 Phlogiston, 125
 Phonetics, 462
 Phonology, 462
 Phonons, 728–729
 Photoelectric effect, 716
 Photons, 679, 737
 Physical chemistry, 88

Physical Cosmology, 199

Physical theory, 673–680

categories of, 675

contemporary, 679–680

explanation in early, 681

foundations of, 673–674

fringe theories of, 675

history of, 675–678

mainstream theories of, 675

modern physics and, 677–678

Newton, classical mechanics, and fundamental forces, 676–677

proposed theories of, 675

theory of everything, 230, 678

Physics

19th century, 707–715

dynamics in, 711–712

elasticity in, 711

heat in, 709

hydrodynamics in, 712–713

key developments in, 708–709

optics in, 713–714

radiation in, 709–710

thermodynamics in, 710

20th century, 735–741

background on, 735–736

constituents of matter in, 735–736

developments in radioactivity and, 736

general relativity in, 325, 678, 739–740

history of the universe and, 740

other significant advances in theory in, 740–741

quantum mechanics in, 737

special relativity in, 738–739, 771

subatomic forces and particle physics in, 737–738

antiquity, 680–686

in Ancient Greece, 681–683

Archimedes in, 682

Aristotle in, 682

explanation in early physical theory in, 681

heavenly bodies and the Earth as round in, 682

Hipparchus in, 682–683

in India and China, 683

role of mathematical theorizing in the West and, 683–686

contemporary, 677–678, 686–692

classical theories in, 688

historical origins of, 687

quantum mechanics in, 689–692

relativity in, 688–689

theory and practice in, 692

gravitational theory, 699–703

cosmology, 701–702

experiment and, 700–701

interpretive questions in, 702–703

language and content of Einstein's theory of, 700

quantum gravity in, 703

quantum theory, 715–721

empiricist interpretations of, 719–720

historical development of, 715–716

mathematical structure of, 716–717

measurement problem in, 718

quantum nonlocality in, 718–719

quantum uncertainty in, 717–718

realist interpretations of, 720–721

wave-particle duality in, 717

solid-state, 726–730

crystals and crystallography in, 727–729

current research topics in, 729–730

high-temperature superconductivity in, 730

quasicrystal or quasiperiodic crystal in, 729

spin glass in, 730

strongly correlated materials in, 729

thermodynamics, 730–735

core concepts in, 731–733

further developments in, 734–735

historical development of, 730–731

laws of, 733–734

Physics, Enlightenment, 692–699

Benjamin Franklin and, 87, 693, 696

Charles Augustin de Coulomb and, 693–694, 696–698, 711

demonstrated through mechanics, 693

Halley's Comet and, 695–696

Joseph Priestley and, 693, 696

science, religion, and, 698–699

significant contributors to, 693–694

thermodynamics in, 694–695

Physics, Middle Ages, 703–707

early medieval society and, 704

expansion of European exploration and thought in, 706

High Middle Ages, 706–707

late Middle Ages, 707

medieval thoughts on space and, 705

in the Middle East, 705–706

philosophers/physicists influential on thought in, 704

Physics, Renaissance, 721–726

astronomy and physics in, 722–726

magnetism in, 726

Physics: the Elements, 543

Physiological chemistry, 66–67, 71

Piaget, Jean, 5, 8, 584, 933

Pickering, Andrew, 838

Pinch, Trevor, 837

Pinchot, Gifford, 243–244

Pinker, Steven, 306, 814

Piso, Willem, 582

Pitts, Walter, 38, 201–203, 207

Plague, the. *See* Black Death

- Planck, Max, 603, 675, 678, 710, 716
 quantum mechanics and, 689, 736
 Planck scale, 603
 Plantinga, Alvin, 498, 945–946
 Plasma membrane, 120–121
 Plasticity, 261
 Plate tectonics, 312–313, 741–747
 constructivist perspective on, 745–746
 continental drift controversy and, 743–746
 emergence of theory of, 742–743
 importance of, 746–747
 rationalist perspective on, 744–745
 Plato, 6, 249, 587, 773, 779, 946
 applicability issue and, 32
 dualism and, 660–661
 geometry and, 316, 318
 on goodness as a form, 590
 idealism of, 434
 informal logic and, 293
 mathematical Platonism and, 784
 Meno, 778–779
 physical theory and, 675
 reasoning and, 783–784
 Platonism, 35, 778–779
 Platykurtic curve, 868
 Pleomorphism, 333
 Pliny the Elder, 550
 PLOR program, 95
 Plotinus, 773
 Podolsky, Boris, 678, 718
 Poincaré, Henri, 175–176, 328–330, 417, 522–523, 712, 735, 875
 Poincaré conjecture, 523, 540–541
 Point-coincidence argument, 702
 Poisons, 579–580
 Poisson, Siméon Denis, 709, 712
 Polányi, Michael, 40, 411, 953–954
 Political authority, 45–46
 Politics of epistemology, 251
 Pollock, John, 945
 Polymerase, 69
 Polymerase chain reaction (PCR), 70
 Polysemy, 822
 Popper, Karl, 57, 217, 220, 222, 283–285, 424, 636, 665, 786, 790
 background of, 284
 critical rationalism and, 900
 falsification model, 220, 283–287, 357–358, 365, 760
 on historicism, 755–756
 induction and, 480
 on inference to the best explanation (IBE), 383
 scientific method and, 285–286
 scientific progress model, 670
 scientific realism and, 799, 801–802
 on scientific revolutions, 809
 theory change and, 902–903
 values in science and, 936
 Population standard deviation, 867
Port-Royal Logic, 785
 Positive feedback, 888
 Positive psychology, 585
 Positivism, 285, 801, 918, 938
 Positron emission tomography, 150
 Possible circumstances, 495
 Possible world semantics, 497
Posterior Analytics, 400–401, 470, 775
 Postmodernist relativism-constructivism, 800
 Postmodern rationality, 774–775
 Postpositivism, 250
 Potential infinity, 416
 Potential outcomes approach (POA), 377–378
 Potential resources, 609
 Pott, Percivall, 562
 Pouillet, Claude, 710
 Pound, Robert V., 701
 Poussin, Charles de la Vallée, 524
 Power and truth, 929
 Power subsets, 826
 Poythress, Vern, 775
 Practical ideal type, 295
 Practical science, 230
 Practice-based accounts, 760
Practice of Geometry, 515
 Pragmatic argument, 754
 Pragmatics, 457, 462, 817
 Pragmatic turn, 634
 Pragmatism, 471, 747–750, 800–801, 928
 democracy and, 749–750
 John Dewey's experimentalism and, 748–749
 pragmatic maxim and theory of truth in, 748
 Richard Rorty's, 749
 Prawitz, Dag, 389
 Praxeology, 249
 Precision, measurement, 16, 20–21
 Precision medicine, 557
 Predicate logic, 292, 497
 Prediction, 296, 750–759
 big data and empirical, 756–757
 empirical evidence, Laplace's demon, and, 754–755
 epistemic and empirical significance of scientific, 754
 explanation and, 668, 751–753
 forecast models and scenarios in, 757–758
 in modeling, 596–597
 probability, Bayes' rule and, 753–754
 in science and society, 758
 Thomas Theorem and self-fulfilling prophecy in, 755–756
 Predictive analytics, 854
 Pre-Euclidean geometry, 317

- Preformationists, 224
 Preparadigm, 635
 Prepared learning, 262
 Presburger arithmetic, 171
Presence of the World, The, 398
 Presentism, 505, 633
 Preservation, 244
 Preservationism, 630–631
 Pressure, gas, 734–735
 Priest, Graham, 631
 Priestly, Joseph, 693, 696
 Prime Number Theorem, 524
 Primitive laws, 447
Prince of Mathematics, 518
 Principal component analysis (PCA), 55
Principia Mathematica, 37, 220, 400, 402–403, 471, 533–534, 642, 851
 Principle of explosion, 498
Principles of Philosophy, 677
 Prior, Arthur, 387
Prior Analytics, 470
 Prior appropriation, 615
 Prisoner's dilemma, 300, 302, 853
 Privacy and big data, 61–62
 Probabilistic explanation, 667
 Probabilistic theory of causation, 554
 Probability, 366, 476, 645, 850, 867–868
 prediction, Bayes' rule and, 753–754
 Problem of circularity, 280
Problems of Philosophy, The, 365
 Procedural knowledge, 435–436
 Process of explanation, 276
 Process ontology, 592
 Process philosophers, 660
 Process-relational paradigm, 592
 Procrustean bed, 850
 Productivity of language, 460
 Progressivism, 110
 Projectability, 480
 Projective geometry, 530
 Prokaryotic cells, 118–119
 Proof and mathematical realism, 782
 Proper function condition and warrant, 945–946
 Proper subsets, 826
 Property dualism, 661
 Propositional logic, 292
 Propositions, 818–819
 Prospect theory, 852
 Protogoras, 784
 Protein Data Bank, 95
 Proteins, 73
 structure and function of, 74
 Protochemistry, 125
 Proto-language, 466
 Protoplasm, 68
 Prototype models, 14
 Prototype theory, 823
 Prout, Marcel, 193–194
 Pseudo-Bayesian approach, 271
 Pseudoenzymes, 75
 Pseudo-patients, 834
 Pseudoscience, 663, 665, 759–763
 critics of distinction of, 759–760
 defenses of distinction of, 760–762
 limits of concept of, 762–763
 multifactorial accounts and, 760–761
 nonscience, bad science, science fraud and, 762
 practice-based accounts and, 760
 practices on the spectrum and, 761–762
 Pseudo-technology, 230
 Pseudotensor, 702
 Psychoanalysis, 814
 Psycho-functionalism, 658
 Psycholinguistics, 146, 458, 463
 Psychological authority, 46–47
 Psychological perspectives of cognitive science, 148
 Psychology, abstraction in, 14
 Ptolemy, 4, 320, 399, 527, 681–683, 704
 Public health, 104–105, 555, 572–573
 Punctualization, 23
 Punctuated equilibrium, 763–766
 origins of theory of, 764–765
 theoretical significance of fossil record and, 763–766
 transformative role of theory of, 765–766
 Pure rationality, 777
 Pure science, 230
 Pure strategy, 299
 Putnam, Hilary, 35, 180, 185, 538
 Putnam, Linda, 777
P value, 352, 356–357, 870
P waves, 311
 Pythagoras, 288, 317–318, 525–526, 674, 683–684, 773
 Pythagoreanism, 32
 Pythagorean theorem, 683

 QBism, 691–692, 719–720
Quadrivium, 515
 Qualia, 661–662
 Qualitative analysis, 29–30
 Qualitative methodology, 156–157, 815
 Qualitative research, generalization in, 306
 Quantificational words, 489–490
 Quantitative analysis, 28–29
 Quantitative information, 397–398
 Galileo's defense of, 400–401
 modern world and, 401–403
 Scientific Revolution as spirit of, 399–400
 Quantitative methodology, 156

- Quantitative research, 812
 generalization in, 306
 Quantum Bayesianism, 691–692, 719–720
 Quantum electrodynamics, 678
 Quantum entanglement, 678
 Quantum gravity, 678–679, 703
 Quantum leap, 689–690
 Quantum mechanics, 689–692, 737
 alternate worlds, 691
 Copenhagen Interpretation, 690–691
 hidden variables, 691
 interpreting, 690–692
 QBism, 691–692
 Quantum nonlocality, 718–719
 Quantum optics, 678
 Quantum physics, 677–678
 Quantum theory, 715–721
 Quartile deviation, 866
 Quasicrystal, 729
 Quasi-experimental design, 255
 Quasiperiodic crystal, 729
 Quate, Calvin, 604
 Quaternions, 521
 Quesada, Francisco Miró, 630
 Quesnay, Francois, 562
Quest for Certainty, The, 748
 Quine, Willard Van Orman, 35, 57, 185, 282, 790, 902
 indispensability and, 781
 metaphysics and, 589–590
 on paradoxes, 641
 Quine-Putnam indispensability arguments, 781
- Rabi, Isidor, 91
 Radiation, 709–710
 Radical doubt, 774
 Radioactive elements, 312
 Radioactivity, 736
 Radiometric age dating, 312
 Radner, Roy, 852
 Ramachandran, G. N., 91
 Ramachandran plot, 91
 Ramanujan, Srinivasa, 532–533
 Ramelli, Agostino, 235
 Ramsden, Jesse, 713
 Ramsey, Frank P., 52, 877
 Ramsey eliminability, 919
 Ramsey sentence of a theory, 52–53
 Ramus, Peter, 784
 Random errors, 19
 Randomization, 267–268
 Randomized controlled trials (RCT), 255, 555
 Random sampling, 868–869
 Rangarajan, Nandhini, 295
 Range, 866
- Rankine, William J. Macquorn, 710, 713
 Raoult, François-Marie, 710
 Rapaport, Anatol, 301, 885
 Rational agents, 232
 Rational choice theory, 299, 854
 Rational expectations equilibrium model, 852
 Rationalism, 424, 774–775, 900
 linguistic, 385–386
 moral, 775
 scientific revolutions and, 812–813
- Rationality, 771–778
 bounded, 777
 coherentism and, 776
 before the common era, 772–773
 evidentialism and, 776
 foundationalism and, 775
 internalism, externalism, and, 776
 and irrationality of groups, 776–777
 perspectivalism and, 775–776
 postmodern, 775–776
 pure, 777
 recent perspectives on, 777–778
 understanding, 772
 Western early modern philosophers on, 773–775
- Rational mechanics, 767–771
 axiomatics and, 770
 history of, 767–769
 mechanicism and, 770–771
 theory of, 769–770
- Ravn, Ole, 323
 Raw data, 214
 Rawls, John, 776
 Raw scores, 867
 R correlation coefficient, 110
 Reaction and neuroscience, 625
 Reaction time, 888
 Reactive oxygen species (ROS), 76
 Realism in mathematics, 778–783
 Benacerraf's problem and, 780–781
 direct realism, 787
 idealism and, 779–780
 indirect realism, 787
 indispensability and, 781–782
 mathematical Platonism, 778–783
 proof, 782
- Realism in science. *See* Scientific realism;
 Scientific realism
- Reasoning, 292–293, 783–788, 905–906
 in the 19th century, 786
 in the 20th century, 786
 in antiquity, 783–784
 deductive, 2, 292–293, 665–668, 787, 905–906
 Enlightenment, 785–786
 forms of, 2, 773–774, 786–788

- history in the West, 783–786
- inductive, 2, 292–293, 498, 665–668, 773–774, 787, 905–906
- mental models and, 584–585
- in the Middle Ages, 784–785
- principles of scientific, 663–665
- process of, 786–788
- speculative, 850–851
- tacit versus explicit, 788
- Rebka, Glen A., Jr., 701
- Received knowledge, 771
- Received view of theories. *See* Theory, syntactic conception
- Recognition and neuroscience, 625
- Recorde, Robert, 528–529
- Recursion, 131–132
- Recursive constructive mathematics, 420
- Recursive function theory, 420
- Redi, Francesco, 332
- Reductionist theory, 276, 557
- Reduction problem, 105
- Reduction sentences, 918
- Reference, 806, 816–818
- Refutation (R), 788–795
 - induction, conjectures, and, 792–793
 - in procedures to confront a theory with evidence, 789–792
- Regiomontanus, 515, 527
- Regnault, Victor, 709
- Regress argument, 426
- Regression analysis, 29
- Regularity accounts, 447–448
- Regularity tradition, 446
- Regulatory control, 105–106
- Reichenbach, Hans, 424, 481, 899
- Reid, Thomas, 924–925
- Reimannian space, 504
- Reinforcing loops, 890
- Relation and neuroscience, 626
- Relative consistency, 795–798
 - applications of, 797
 - completeness and, 795
 - incompleteness and, 795–796
 - value of, 796–797
- Relative dating, 312
- Relative scarcity, 611
- Relativism, linguistic, 460–461
- Relativistic standard cosmology, 198–199
- Relativity. *See* General relativity; Special relativity
- Relevant logics, 631
- Reliabilism, 250
- Reliability
 - in conceptual analysis, 188
 - evidence, 256–257
- Religion and science
 - during the Enlightenment, 698–699
 - during the Renaissance, 578–579
- Religious authority, 46
- Remembering and neuroscience, 625
- Renaissance, the
 - See also* Enlightenment, the
 - historical changes and epistemological shifts during, 809–811
 - mathematics of, 525–532
 - medicine in, 577–583
 - physics of, 721–726
- Renewable resources, 609
- Rényi information measures, 397
- Replication, experiment, 20
- Representation, measurement as, 543–544
- Representationalism, 383–384
 - structuralism and, 876–878
- Representational theory of measurement, 544
- Representative values, 864
- Representing and Intervening*, 904
- Reproduction and evolutionary psychology, 263–264
- Reproductive value, 264
- Research hypothesis, 870
- Research questions, 906–908
- Res extensa, 774
- Resistant statistic, 866
- Res logicans, 774
- Resource exploration, 310
- Responding to emotions in patient-centered communication, 345–346
- Response surface, 269–270
- Ressel, Josef, 713
- Retrieval, data, 211
- Retroduction, 803
- Retrotransposons, 85
- Reyna, Valerie, 144
- Rhazes, 559
- Rheology, 310–311
- Rheticus, Georg Joachim, 529
- Rhetoric, 784
- Ribet, Ken, 540
- Ribosome, 121
- Richards, I. A., 181
- Richarz, Franz, 712
- Richman, Fred, 420
- Riemann, Bernhard, 325, 328–330, 519–520, 673
- Riemann hypothesis, 520
- Riemannian space, 519
- Riemann surfaces, 519
- Riemann zeta function, 520
- Riese, Adam, 528
- Rigid bodies, 769
- Riparian law, 615

- Risk science, 852
 Rivlin, Ronald, 769
 RNA (ribonucleic acid), 73, 78, 82, 93
 Robert, Julius, 708
 Robinson, Julia, 522, 538–539
 Robinson hypothesis, 538
 Robust designs, 272–273
 Rocks and minerals, 313–315
 Roebuck, John, 125
 Rogers, Carl R., 427–428, 933
 Rohrer, Heinrich, 604
 Röntgen, Wilhelm Conrad, 89, 736
 Root mean squared error of approximation (RMSEA), 601
 Rorty, Richard, 749, 928, 954–955
 Rosen, George, 550
 Rosen, Nathan, 678, 718
 Rosenberg, Charles, 573
 Rosenhan, David, 834
 Rothmund, Paul W. K., 605
 Rothman, Kenneth, 555
 Routley, Richard, 631
 Roux, Wilhelm, 941
 Rowland, Frank S., 128
 Rowland, Henry A., 712, 714
 R square, 597
 Rubik, Ernő, 540
 Rubik's Cube, 540
 Rudbeck, Olaus, 561
 Rudolff, Christoph, 528
 Ruesch, Jürgen, 159–160
Rule of Descartes, The, 507
 Rule of succession, 290
 Ruse, Michael, 285
 Ruska, E., 603
 Russell, Bertrand, 37, 219–220, 293, 336, 365, 417, 481, 534, 928
 justification of induction, 481
 mathematical logic, 470–471
 neutral monism and, 660
 phenomenalism and, 655
 rationality and, 774, 776
 Russell's paradox, 219–220, 640–642, 830
 Russell, John, 708
 Russell, Scott, 713
 Russell, Stuart, 39
 Russell's paradox, 219–220, 640–642, 830
 Rynasiewicz, Robert, 919

 Saccheri, Gerolamo, 325–326
 Sackett, David, 555
 Sackur-Tetrode equation, 432
 Safety factors, 233–234
 Salam, Abdus, 675, 738
 Salmon, Wesley, 275–276, 368
 Salumbrino, Agostino, 580
 Sample standard deviation, 867
 Sampling distribution, 869
 Sampling variability, 869
 Samuel, Arthur, 178
 Sapir, Edward, 461
 Sapir-Whorf hypothesis, 461
 Sauce béarnaise effect, 262
 Saussure, Ferdinand de, 183, 461, 464
 Scale separation, 98–99
 Scarcity, 611–612
 Schaeffer, Francis, 951
 Scheffé, Henry, 872
 Schelling, F. W. J., 951
 Schelling, Thomas, 301
 Schema, 463
 Schema-abstraction, 14
 Schemas, cognitive science, 144–145
 Schematic construct, 462
 Schiller, Friedrich, 30
 Schleiden, Matthias Jakob, 116–117, 565
 Schleiermacher, Friedrich, 951
 Schlick, Moritz, 918
 Scholastic philosophy, 559
 Scholes, Myron, 852
 Schönflies, Aruthur, 89
 Schopenhauer, Arthur, 181
 Schotch, Peter, 630
 Schrieffer, Robert, 686
 Schröder, Ernst, 219, 472
 Schrödinger, Ernst, 105, 678, 690, 716, 737
 Schrödinger's Cat, 678
 Schutz, Alfred, 154
 Schwann, Theodor, 116–117, 565
 Schwartz, David, 93
 Schwarzschild, Karl, 330
 Schwinger, Julian, 678
 Science
 ambivalence of, 832–833
 applicability of mathematics in, 31–35
 paradigms in development of, 635–636
 and religion in the Enlightenment, 698–699
 theory of, 637
Science, the Endless Frontier, 127
Science, Technology, and Society in 17th Century England, 937
 Science and Technology Studies, 836
 Science fraud, 762
Science of Mechanics, The, 771
 Scientific abstraction, 12–13
 Scientific animation, 454–455
 Scientific appropriations of worldview, 953–954
 Scientific authority, 45, 672–673
 Scientific community, 725

- Scientific computing, 289
- Scientific determinism, 755
- Scientific framework of syntax, 882–883
- Scientific Image, The*, 405
- Scientific management, 238
- Scientific method, 294
 - applied to complementary medicine, 167–168
 - consilience as, 192–193
 - Scientific Revolution and, 811–812
- Scientific physiology, 559–560
- Scientific realism (SR), 12, 279–281, 799–807, 920
 - epistemological theses in, 805–806
 - instrumentalism and, 404–405
 - methodological theses in, 807
 - naive realism, critical realism, and, 802–804
 - nomological, 446–447
 - ontological theses in, 804–805
 - and other schools of the philosophy of science, 800–802
 - philosophy of science in, 804
 - quantum theory and, 720–721
 - realism and antirealism and, 799–800
 - semantic theses in, 806–807
 - structuralism and, 875–876
- Scientific reasoning, principles of, 663–665
- Scientific revolutions, 668–671, 785–786, 807–816, 903, 905
 - 20th-century conceptual revolutions, 813–815
 - emergence of, 808–809
 - historical changes and epistemological shifts during the Renaissance and, 809–811
 - philosophy of science, rationalism, and enlightenment and, 812–813
 - Renaissance mathematics and, 529–530
 - scientific method and, 811–812
- Scope conditions, 913
- Scripta*, 527
- Seafloor spreading, 743
- Searchlight theory of science, 900, 936
- Searle, John, 39, 817
- Second-order cybernetics, 202–203
- Second-order logic, 512
- Sedimentary rocks, 314
- Sedol, Lee, 41
- Seismic waves, 311
- Selection, 308–309
- Seleucus of Seleucia, 682
- Self-efficacy, 436
- Self-fulfilling prophecy, 755–756
- Selfish Gene, The*, 84, 619
- Self-organization in complex systems, 174–175, 887
- Self-reference in cybernetics, 204–206
- Self-reflection, 251–252
- Selfridge, Oliver, 36
- Self-verification, 790–791
- Seligman, Martin, 585
- Selten, Reinhard, 301
- Selz, Otto, 284
- Semantic ambiguity, 822
- Semantic analysis, 795
- Semantic empiricism, 409
- Semanticity in language, 460
- Semantic Leaps: Frame-Shifting and Conceptual Blending in Meaning Construction*, 191
- Semantic realism, 806–807
- Semantic relations, 183
- Semantics, 462
 - of interpretation, 406–410
 - as linguistic framework, 457
- Semantics (introduction to theory), 816–820
 - communicative meaning and, 817–818
 - descriptive meaning of sentences and, 818–819
 - of grammatical forms, 818
 - linguistic approach to, 819
 - making inferences, 817
 - meanings as concepts and, 818
 - semiotic triangle, 819
 - sentence meaning and compositionality and, 818
 - truth condition and, 816–817
- Semantics (scientific and empirical), 820–824
 - expressive meaning, 822–823
 - homonymy, 821–822
 - lexemes and lexicon of a language and, 821
 - meaning and cognition, 823
 - polysemy, 822
 - semantic ambiguity or vagueness, 822
 - synonymy, 822
- Semantic subset principle, 487
- Semantic theory, 388–389
- Semantic theses in scientific realism (SR), 806–807
- Semiotic triangle, 181, 819
- Semmelweis, Ignaz, 370, 573
- Sense of the game, 411
- Sensing plus sensemaking, intuition as, 414–415
- Sensory communication, 92
- Sentence analysis, 879–880
- Sentence meaning, 818
- Sentence parsing, 879
- Sentences, descriptive meaning of, 818–819
- Sentence semantics, 818
- Sentential logic, 292
- Separation, axiom schema of, 50–51
- Sequential reasoning, interplay between intuition and, 414
- Serre, Jean-Pierre, 537, 540
- Servetus, Michael, 580
- Set theory, 824–831
 - as formal language, 825–827
 - as foundation of mathematics, 830–831
 - as science of infinity, 827–830

- Seven Bridges of Königsberg problem, 509, 523
 Sex differences in the ideal mate, 263–264
 Sexual selection, 260–261
 Shadow zone, 742
 Shaffer, Simon, 838
 Shannon, Claude, 152, 158–159, 202, 290, 520
 Shannon information measure, 394–397
 Shapin, Steven, 838
 Shapley, Harlow, 196
 Shaw, Cliff, 36
 Shaw, Donald, 187
 Sheaf theory, 537
 Shen Kuo, 683
 Shields, Patricia, 295
 Shimura-Taniyama-Weil theory, 537
 Shimure, Goro, 540
 Shot noise, 211
 Signal theory, 151–152
 Significant figures, 20
Silent Spring, 128
 Siliciclastic rocks, 314
 Silverman, Daniel, 76
 Silvester, Trevor, 148
 Simmonds, James G., 644
 Simon, Herbert, 36–37, 777, 853, 919
 Simple regularity view of scientific laws, 439–440
 Simple systems, 887
 Simpson, George Gaylord, 764
 Simulated annealing, 95
 Single-photon emission computed tomography (SPECT), 626
 Singular perturbations, 650–652
 Situational and self-awareness, 347
 Sixth sense, 412
 Sjøberg, Dag I. K., 844
 Skepticism, 434
 climate, 142–143
 organized, 936
 Skewness, 868
 Skolem, Thoralf, 219
 Skovmose, Ole, 323
 Skyrms, Brian, 946
 Slipper, Vesto M., 195
 Slow life history strategy, 261
 Smalley, Richard W. E., 128, 604
 Smallpox, 582
 Smeaton, John, 236, 694
 Smith, Adam, 243, 610–611
 Sneed, J. D., 53, 919
 Snell's law, 677
 Snow, Charles Percival, 192–193
 Snow, John, 371
 Sobel, Michael, 601
 Social construction of scientific knowledge, 831–840, 938–939
 ambivalence of science and, 832–833
 anthropological and ethnomethodological investigations in, 837–838
 controversy and future directions in, 838–839
 criticisms of, 838
 ethnographic laboratory studies in, 836–837
 Gestalt shifts and incommensurability in, 834–835
 incorporating nonhuman actants in, 838
 paradigms as interpretive constraints, 834
 Robert Merton and institutional imperatives in, 832
 thought collectives and paradigms in, 833–834
 trading zones and boundary objects in, 835–836
 Social contract theory, 774
 Social exchange theory, 393
 Social forecasting, 854
 Social health, 550
 Social network analysis, 144
 Social relations and evolutionary psychology, 265–266
 Social systems model, 889
 Sociobiology, 194
 Socioenvironmental theories, 348
 Sociolinguistics, 146, 458, 463
 Sociological perspectives of cognitive science, 148
 Sociology of scientific knowledge (SSK), 22, 903
 Socrates, 434, 505, 773, 784
 Socratic dialogues, 773
 Soft systems methodology (SSM), 173
 Software engineering, 840–846
 definition and function of software and, 841
 as a discipline, 841–842
 engineering as concept and profession and, 841
 process in, 842–843
 theoretical and transdisciplinary foundations of, 843–844
 theory use and development in, 844–845
 See also Computer science
 Soil science, 846
 Sokal, Alan, 838–839
 Solar system
 antiquity physics and, 682
 Isaac Newton on, 321–322, 325, 699
 LaPlace and, 510–511
 medieval thoughts on, 705
 Solid-state physics, 726–730
 Soliton, 708
 Soma Cube, 540
 Somayahi, Nilakantha, 683
 Sommerfeld, Arnold, 89
 Sophie's choice, 850
 Sophisticated regularity view of scientific laws, 440–441
 Sorensen, R., 922
 Sorites paradox, 639–640, 642
 Sosnoski, James, 191

- Sound change in language, 468–469
- Soundness, 846–850
 - and completeness in mathematical logic, 849
 - in deductive logic, 846–847
 - in formal or symbolic logic, 848–849
 - logical form and, 847–848
- Source rate, information, 290
- Sowa, John, 181
- Spallanzani, Lazzaro, 450
- Spanish Flu of 1918, 369
- Special efficacy of measurement, 546
- Special relativity, 688–689, 738–739, 771
- Speculation, 850–855
 - in development of mathematical models, 851–852
 - in development of nonrational models, 852
 - economic, 852–853
 - as forecasting, 854
 - in game theory and decision theory, 853–854
 - in models of speculative reasoning, 850–851
- Speech act, 462
- Spencer, Herbert, 196, 434
- Spengler, Oswald, 690
- Spin control, 25
- Spin glass, 730
- Spinoza, Baruch de, 323, 774–775
 - neutral monism and, 660
 - reasoning and, 785
- Spin-statistics theorem, 737
- Spirantization, 468
- Squaring a circle, 685–686
- Stahl, George Ernst, 125, 548, 564
- Standard deviation (SD), 867–868
- Standard error of the mean (SE_m), 869
- Standard gel electrophoresis, 93–94
- Standardized root mean square residual, 601
- Standard score, 867
- Standard uncertainty, 19
- Stanley, Wendell Meredith, 69
- Star, Susan Leigh, 835
- Statement view, 912
- Statics, 769
- Statistical generalization, 306
- Statistical inference, 366–367, 377–378
 - frequentist, 858–863
 - frequentist error probabilities, 861–862
 - inferential claims for estimation and testing, 862
 - model-based, 858–863
 - reliable inference and trustworthy evidence, 862–863
- Statistical inference, Bayesian, 377–378, 855–858
 - applications of, 857–858
 - challenges to, 857
 - hypothesis tests and Bayes factors, 856–857
- Statistical laws, 449–450
- Statistical linguistics, 458, 463
- Statistical mechanics, 429–430
- Statistical relevance (S-R) model, 368
- Statistical testing, 869–870
- Statistics, 863–870
 - abduction in, 3
 - Bayesian, 356, 377–378
 - completeness in, 870–874
 - complete sufficiency and similar tests and, 873
 - modeling versus inference, 873–874
 - sufficiency, ancillarity, and, 872–873
 - data and data sets, 864
 - descriptive, 864–867
 - frequentist, 377
 - inferential, 864, 867–868
 - in mental models, 585–586
 - random sampling, 868–869
 - statistical testing, 869–870
- Statistics and probability, 289–290
- Steady state, 197–198
- Stebbins, G. Ledyard, 83
- Stefan, Josef, 709
- Stereochemistry, 89
- Stevens, Stanley, 543
- Stewart, Balfour, 709
- Stich, Stephen, 185
- Stifel, Michael, 528
- Stinchcombe, Arthur, 359–360
- Stochastic processes, 851
- Stoicism, 365
- Stokes, George, 713
- Stomatolites, 316
- Stone, Marshall, 219
- Stout, Alicia, 538
- Strachey, Christopher, 178
- Strange attractors, 539
- Strategic rationality, 301
- Stratisgraphy, 312
- Strawson, Peter, 481
- Street, Richard L., 344–346
- Strict implication, 630–631
- String theory, 533, 536, 675, 678, 680, 686–687, 702
 - super, 702
- Strong inductive arguments, 494–495
- Strong programmes, 836
- Structural biology, 94–95
- Structural empiricists, 876
- Structural equation modeling (SEM), 599–601
- Structural geology, 310
- Structuralism, 445, 874–878
 - in the scientific realism debate, 876–878
 - in the scientific theory and representation debates, 876–878
- Structural modeling approach, 378

- Structure and Confirmation of Evolutionary Theory, The*, 916
- Structure of Scientific Revolutions, The*, 45, 219, 286, 663, 808, 833, 901, 938, 953
- Struggle for life, 302
- Strutt, John William, 713
- Subatomic forces, 737–738
- Subduction, 313
- Subgame perfect equilibrium, 301
- Subjective idealism, 654
- Subjective probabilities, 301, 753–754
- Subjectivism, 251
- Subjectivity of abduction, 3–4
- Sublimation, 791
- Subordination, 880
- Subsets, 826–827
- Substance dualism, 661
- Substance ontologies, 592
- Substitute models, 915
- Sudoku, 540
- Suess, Eduard, 743
- Sufficiency-necessary theory, 554
- Sufficient-component theory of causation, 554–555
- Sugar metabolism and enzyme specificity, 66
- Sulfuric acid, 125
- Summa de Arithmetica, Geometrica, and Proportioni et Proportionalita*, 527–528
- Summary statistics, 864
- Sumner, James B., 69
- Sum of squares (SS), 867
- Supernatural, the, 412
- Super-Nostratic language, 466
- Supervenience, 773
- Suppes, Patrick, 213–214, 816
- Surgery, 561–563
- Surreal numbers, 540
- Survival and evolutionary psychology, 262–263
- Sustainability, 246–247
- S waves, 311
- Swineshead, Richard, 707
- Sydenham, Thomas, 558–559, 563
- Syllables, 462
- Syllogisms, 217, 433, 470
- Aristotelian, 217–219, 291, 663
- Sylvius, Jacobus, 560
- Symbiogenesis, 83
- Symbolic expert systems, 37–38
- Symbolic generalizations, 634
- Symbolic interactionism, 250–251
- Symbolic logic, 492, 497
- soundness in, 848–849
- Symbolic reasoning systems, 36–37
- Synapses, 71
- Synchorinic meaning, 465
- Synchrony, 461
- Synonymy, 822
- Synopsis of the Astronomy of Comets*, 695
- Syntactic analysis, 879–880
- Syntactic concept of theory, 817
- Syntactic gradience, 882
- Syntactic method, 795
- Syntactic theories of interpretation, 408–409
- Syntactic typology, 881
- Syntagmatic relations, 183
- Syntax, 462, 912
- Syntax (introduction to theory), 878–882
- approaches to syntactic analysis and, 878–880
- cohesion and coherence, 880
- sentence analysis through immediate constituents and parsing, 879–880
- subordination and coordination, 880
- Syntax (scientific and empirical), 881–884
- clause grammar and structure, 881–882
- conditional syntactic structures, 882
- empirical framework of, 881
- government and binding, 883
- of negation, 883–884
- scientific framework of, 882–883
- statistical analysis, 884
- syntactic gradience and fuzzy syntactic structures, 882
- syntactic typology, 881
- transformational and generative grammar, 883
- Synthesis, 435
- intuition and, 412
- System, measurement, 18–19
- Systema Naturae*, 563
- Systematic errors, 19, 267–268
- Systematic Treatise on the Principal Disease of the Interior Valley of North America*, 549
- System dynamics model, 889–890
- Systemic perspectives in cognitive science, 144
- System of Logic*, 32, 363–364, 786
- System of the World, The*, 511
- Systems biology, 100–101
- Systems concepts, 886–888
- Systems models, 888–890
- causal relations model, 890
- complex adaptive systems (CAS), 173, 175, 889
- social, 889
- system dynamics, 889–890
- viable, 890
- Systems neuroscience, 623
- Systems science, 884–891
- systems concepts in, 886–888
- systems models in, 888–890
- systems thinking in, 885–886
- Systems theory, 176, 392
- Systems thinking, 885–886

- Tabula rasa, 774
 Tabula rasa, 434
 Tacit versus explicit reasoning, 788
 Taguchi, Genichi, 267, 272
 Tait, Peter, 712
 Tamm, Igor, 729
 Taniguchi, Norio, 604
 Taniyama, Yutaka, 540
 Tantia, Jennifer Frank, 412
 Tarski, Alfred, 51, 219, 641, 817, 914, 929
 reformulation of correspondence theory of truth, 793
 Tartaglia, Niccolo Fontana, 528
 Tatum, Edward, 82
 Taxonomy, 893–897
 biological, 894–895
 as conceptual categories, 893
 as controlled vocabulary, 894
 examples of taxonomic names, 895–896
 international and universal coding systems for, 895
 other applications of, 896
 process of taxonomic classification, 895
 Taylor, Charles, 951, 955
 Taylor, Frederick Winslow, 238
 Taylor, Joseph H., Jr., 701
 Taylor's theorem, 647
 Taylor-Wegener hypothesis, 745
 Teal, Cayla R., 346
 Team mental models (TMMs), 585
 Technological perspectives on cognitive science, 149
 Technology
 engineering theory and, 231–232
 engineers and philosophy of, 232–233
 Technology doping, 606
 Teilhard de Chardin, Pierre, 30
 Telford, Thomas, 236
 Temporal precedence, 256
 Tennant, Niel, 389
 Ten Rhijne, Willem, 582
 Tense forms, 469
 Tesla, Nikola, 40, 236
 Tests of significance, 869–870
 Tetlock, Phil, 854
 Thagard, Paul, 904
 Thales of Miletus, 317, 505, 525–526, 681, 697, 772–773
 Theolinguistics, 464
 Theoretical framework, 294
Theoricae Novae Planetarum (New Theories of the Planet), 527
Theorie analytique de la Chaleur, 512
 Theory, 153
 analytic, 153
 as axiom, 910–911
 of causation, 554–555
 classical view of, 816
 completeness of, 170–171
 critical, 154
 in engineering, 233–234
 evaluated using metatheory, 593–594
 evidence in support of, 257–258
 experiential, 156
 experimental, 154, 156
 metatheory behind, 592–593
 as model, 911–912
 philosophy of science and, 663–665
 rational mechanics, 769–770
 of science, 637
 semantic conception, 913–917
 advantages of using semantic view, 915–916
 classical view of theories, 914
 construction of theoretic models, 915
 foundation of, 914–915
 laws of nature and, 916–917
 role of semantic models in science, 915
 in software engineering, 844–845
 syntactic conception, 409, 816, 917–920
 demise of syntactic view, 920
 eliminability of theoretical terms, 918–920
 empirical significance of theoretical terms, 918
 formative influences, 917–918
 partial structures and empirical content, 919–920
 Ramsey eliminability and empirical content, 919
Theory Building, 364
 Theory change, 897–904
 from logical analysis to historical change, 898–901
 as practice, 903–904
 as radical shift, 901–903
 Theory construction, 817, 905–909
 forms of reasoning and, 905–906
 hypotheses and research questions in, 906–908
 professional scrutiny in, 907–908
 salience of, 908–909
 theory origination and, 905
 Theory-ladenness of observation, 902
 Theory of a case, 231
 Theory of affordances, 393
 Theory of argumentation, 426–428
 Theory of everything, 230, 678
 Theory of fluxions, 531
 Theory of forms, 773
Theory of Games and Economic Behavior, 299
 Theory of J, 423–426
 Theory of medical knowledge and evidence, 556
 Theory of relativity, 325, 677, 771
 See also Einstein, Albert
 Theory of schemes, 289, 537
 Theory of topoi, 537

- Theory structure, 909–913
 additional dimensions of, 912–913
 theories as axioms and, 910–911
 theories as models and, 911–912
 Thermal cyler, 70
 Thermodynamics, 89, 195–196, 710, 730–735
 core concepts of, 731–733
 in Enlightenment physics, 694–695
 historical development of, 730–731
 kinetic (molecular) theory and, 429–430
 laws of, 733–734
Thesaurus Geographicus, 745
 Thesis of structural identity, 752
 Theta theory, 883
 Thinking
 abstract, 6–10
 by artificial intelligence (AI), 38–39
 Thinking aloud technique, 36
 Thinking body, 774
 Third dogma of empiricism, 634
 Third realm, 779–780
 Thomas, Brian, 96
 Thomas theorem, 755–756
 Thompson, Benjamin, 710
 Thomson, J. J., 127, 736
 Thomson, William, 731
 Thought collectives, 833–834
 Thought experiments, scientific and philosophical, 920–927
 Charles Darwin's *Origin*, 923
 further reflection on nature of, 925–926
 Galileo Galilei and falling bodies, 922
 John Locke's account of personal identity, 924–925
 nature and types of, 921–922
 Thomas Hobbes and state of nature, 923–924
 Thought trials, 905
 Three-body-problem, 522–523
 Threshold-dependent noncooperative behavior, 302–303
 Tibshirani, Robert, 597
 Tiles, Mary, 416
 Tirole, Jean, 45
 Tit-for-tat strategy, 299, 301–302
 Tomasello, Michael, 306
 Torricelli, Evangelista, 768
 Torsion balance, 697
 Total orbs, 705
 Toulmin, Stephen, 221, 280, 283, 427, 786
 Trading zones, 835–836
 Traditional authority, 46
 Traditional Chinese Medicine, 550–551
 Tragedy of the Commons, 244–245, 613–614, 854
Traité de Dynamique, 693–694
 Transcendental idealism, 32, 434, 654
 Transcendentalism, 800
 Transcranial Doppler (TCD) ultrasound, 627
 Transcranial magnetic stimulation, 150
 Transformational grammar, 883
 Transform boundaries, 313
 Translation (proteins), 74
 Translation and cognitive science, 146
 Travers, Morris William, 710
 Trawick, Sharon, 837
Treatise on Electricity and Magnetism, 90
Treatise on Philosophy, 772
 Trigonometry, 529
 Trimmed means, 866
 Tripartite theory of knowledge, 249
 Trisecting the angle, 539
 Trivers, Robert, 263
 Tropes of speech, 151
 Tropical illness, 549, 575
 True contradiction, 498
 True generalization, 440
 Trueness, 16
 Truesdell, Clifford Ambrose, III, 769
 Truth, 472, 927–929
 coherence theory, 927–928
 correspondence theory, 927
 explanation and, 273–274
 Foucault, power, and, 929
 minimalism and deflationism and, 928–929
 pragmatism and, 928
 rationality and, 776
 synthetic theory of, 806
 Truth conditions, 156, 819
 denotations and, 819
 semantics and, 816–817
 Truth-functional logic, 497
 Truth in advertising, 436
 Truthiness, 776
 Truth regimes, 929
 Truth-value judgment task (TVJT), 487–488
 T-schema, 51
 Tuberculosis, 548
 Tucker, Albert W., 300
 Tucker Lewis index, 601
 Tukey, John W., 841
 Tulp, Nicholas, 561
Turin Academy, 510
 Turing, Alan, 38–39, 130, 178, 290, 393, 420, 535–536, 658, 786
 Turing computable, 131–134
 Turing machine, 132–134, 178, 393, 397, 536
 Turing test, 536, 658
 Turner, Jonathan, 911
 Turner, Mark, 190
 Tversky, Amos, 412, 476, 852

- Twentieth century
 - the, biochemistry in, 68–71
 - biophysics in, 90–97
 - cybernetics in, 201–208
 - mathematics in, 532–541
 - physics in, 735–741
 - reasoning in, 786
- Two New Sciences*, 724
- Two-tail test, 352–354
- Type A error, 19
- Type B error, 19
- Type I error, 110, 354–356
- Type II error, 355–356
- Type III error, 355–356
- Typological reconstruction, 465–466
- Ultramafic magma, 314
- Uncertainty management in patient-centered communication, 346
- Underdetermination of theory by data, 902
- Understanding, 931–934
 - developmental stage of, 933
 - graduated, 931–933
 - personal, 933–934
- Unified field theory, 678
- Uniformity of nature, 753
- Unimodal distribution, 866
- United Protein Database, 78
- Universal accounts, 446–447
- Universal characteristic, 219
- Universal grammar, 472, 485
- Universalism, 761, 936
- Universal statements, 286
- Universe, history of the, 740
- Unreasonable effectiveness of mathematics, 781–782
- Unstandardized scores, 867
- Urdu, Mo'ayyeddin, 705
- Urea, 126
- Uses and gratifications theory, 392
- U.S. National Center for Biotechnology Information, 78
- Ussing, Hans, 94
- Utterance meaning, 816
- Vaccine research, 370
- Vacuoles, 121
- Valence, 183
- Valid deductive arguments, 493–494
- Validity
 - in conceptual analysis, 188–189
 - evidence, 256–257
- Value-free science, 672–673
- Value of logic, 498
- Values in science, 935–940
 - constitutive values of science and, 935–937
 - direction of science and, 937–938
 - place of extrascientific, 937–939
 - scientific worldview and, 938–939
 - value-free science and scientific authority and, 672–673
- Van der Waals, Johannes Diderik, 710
- Van der Walt, Christa, 343
- Van de Vijver, Fons J. R., 193
- Van Fraassen, Bas, 405, 816–817, 914–916
- Van Leeuwenhoek, Anton, 116–117
- Van Rensburg, Gisela, 343
- Van't Hoff, Jacobus Hendricus, 89, 126–127
- Variability
 - measures of, 866–867
 - sampling, 869
- Variables, 864
- Variance, 867
- Variation and fuzziness in language, 467–468
- Vasiliev, Nicolai A., 629
- Vaughan, Frances, 413–414
- Vauquelin, Nicolas Louis, 64
- Vector space model (VSM), 182
- Velarization, 468
- Venn, John, 524, 786
- Venn diagrams, 524
- Verb-object (VO) language, 466
- Verisimilitude, 58, 793–794
- Vermuyden, Cornelius, 235
- Vesalius, Andreas, 558, 560–561
- Vespucci, Amerigo, 528
- Viability, 890
- Viable systems model, 890
- Vienna Circle, 806, 899, 917–918
- Virchow, Rudolf, 116–117, 565
- Virtual bosons, 679
- Viruses, 122
- Visceral sensations, 414–415
- Visualization, 29
 - data, 215
- Visual sieve, 539
- Vitalism, 940–944
 - evolution and, 942–943
 - Gaia hypothesis and, 943
 - Hans Dreisch and, 940–941
 - and organicism in the 20th century, 941–942
- Vitamins, 70
- Vitruvius, 319–320
- Vocabulary change in language, 469–470
- Voice of medicine, 344
- Voice of the lifeworld, 344
- Volta, Alessandra, 87, 125, 697, 708
- Voltaic pile, 708
- Voltaire, 110, 506

- Volterra, Vito, 34
 Von Bekesy, Georg, 92
 Von Haller, Albrecht, 87, 564
 Von Hohenheim, Theophrastus Bombastus, 560
 Von Jacobi, Moritz, 712
 Von Jolly, Philipp, 712
 Von Mayer, Julius Robert, 88, 710, 731
 Von Neumann, John, 289, 299–301, 536, 716, 853
 Von Peuerbach, Georg, 527
 Von Suppé, Franz, 817, 915
 Vygotsky, Lev, 8, 434, 933

 Waage, Peter, 88, 710
 Waddington, C. H., 940
 Waldegrave, James, 299–300
 Wallace, Alfred R., 80, 616
 Wallis, John, 507
 Wallis, Steven E., 594
 Walsh, Denis, 943
 Walsh, John, 87
 Walton, Douglas, 786
 Wang, Yingxu, 841
 War injuries and trauma, 580
 Warrant, 945–948
 Gettier problem, proper function condition, and, 945–946
 justification
 foundationalism, and coherentism in, 947–948
 internalism, and externalism in, 946–947
 Watson, James, 73, 82, 93, 96
 Watson, John B., 225
 Watt, James, 236
 Watts, Michelle, 76
 Wave-particle duality, 717
Way We Think: Conceptual Blendin and the Mind's Hidden Complexities, The, 190
 Weak evidence, 253–254
 Weak inductive arguments, 494–495
 Weber, Heinrich Friedrich, 709
 Weber, Max, 46, 776, 937
 Weber, Wilhelm, 519
 Wegener, Alfred, 741–747
 Weil, Andre, 289, 536–537
 Weil representation, 537
 Weinberg, Jonathan, 185
 Weinberg, Steven, 702, 738
 Weinberg, Wilhelm, 451
 Weiss, Paul, 176
 Weldon, W. F., 110
 Well-formed formulas (WFFs), 471
Weltanschauung. See Worldview
 Werner, Johannes, 529
 West, Jennifer, 605
 Weyl, Hermann, 420
Whetstone of Witte, The, 528
 Whetten, David, 906
 Whewell, William, 192–193, 788
 Whiggism, 633
 Whig interpretation of history, 633
 Whipple, Fred, 695
 Whitehead, Alfred North, 37, 176, 220, 336, 534, 642
 mathematical logic, 470–471
 metaphysics and, 590
 panpsychism and, 660
 Whittaker, Robert H., 895
 Whole-part fallacies, 476
 Whorf, Benjamin, 7–8, 461
 Widman, Johann, 528
 Wieck, Karl, 905
 Wiedemann, Gustav, 709
 Wien, Wilhelm, 736
 Wiener, Norbert, 152, 177, 201–202
 Wigner, Eugene, 33, 781
 Wiles, Andrew, 539
 Wilkins, Maurice, 93
 Will, Frederick L., 481–482
 William of Ockham, 12–13, 15, 559, 589, 706, 850
 Williams, Donald, 590
 Willstätter, Richard, 69
 Wilson, Edward O., 192–194
 Wilson, J. Tuzo, 312
 Winfree, Erik, 605
 Winsorizing, 866
 Wittgenstein, Ludwig, 221, 293, 665, 901
 on philosophical propositions, 789
 Woese, Carl, 84, 895
 Wöhler, Friedrich, 126
 Wolf, Karl, 363–364
 Wolff, Caspar F., 564
 Wolff, Johann Friedrich, 564
 Wollaston, William Hyde, 713
 Wolters, Albert, 951
 Women's health care, 348–349
Women's Ways of Knowing, 934
 Woodward, James, 209–212
 Woodward, Robert Burns, 127
 Woolgar, Steve, 837
 WordNet, 181
 Word webs, 586
 Work-energy principle, 769
 Working hypotheses, 295
World Hypotheses: A Study in Evidence, 592
 Worldly structure, 408
 Worldview, 948–956
 critiques of, 955–956
 history of, 950–951
 linguistic appropriations, 954–955
 in mental models, 585

- phenomenological appropriations, 952
- salience of, 948
- scientific appropriations, 953–954
- structure of, 948–950
- uses of, 951–955
- values in science and, 938–939
- World War II, engineering through, 237–238
- World Wide Web (WWW), 78
- Wright, Orville, 238
- Wright, Sewall, 81
- Wright, Wilbur, 238

- Xenophanes, 286
- X-ray computed tomography, 150
- X-ray crystallography, 88–89, 91
- X-ray spectrogoniometer, 91

- Yang, Ning, 686–687
- Yang Hui, 514
- Yang-Mills theory, 686–687
- Yin, Robert K., 306

- Yin and yang, 550–551, 568–569
- Young, Charles, 187
- Young, Thomas, 708, 714
- Yuelin, Jin, 481

- Zadeh, Lotfi A., 178–179
- Zeno, 526
- Zeno's paradox, 639, 642
- Zermelo, Ernst, 299–300, 330, 538, 830
- Zermelo-Fraenkel theory, 538, 830
- Zeroth law, 734
- Zewail, Ahmed, 128
- Zhou Daguan, 569
- Zilsel, Edgar, 634
- Zimbardo, Philip, 777
- Zimmermann, Eberhard A. W., 563
- Zone of proximal development (ZPD), 933
- Zoonomia*, 80–81
- Z-score, 867
- Zuckerberg, Mark, 239
- Zymase, 69

